

Figure 1

CGAGAGACGACAGAAGGTTACGGCTGCGAGAAGACGACAGAAGGGTCCAGAAAAA  
GGAAAGTGCTGGAGGGGAGTGGGGACAAAAGCAGCGACCAAGTGAATGTCACTTC  
AGTGACTGAGGCCAGGCAAAACGCGCGGGAAGGATTTTGTGTAGCTTGGGACCTT  
TCATAGACACTGATGACACGTTTACGCAAAATAGAAATTTGAGGAGAAACGCCTGG  
GCCTTCGGAAAGGAGTGATTGATTAGTACTTGCAAGTTTAGGTGACTTTAAGGAGAA  
CTAACTAATGTATACTATTGAGGGAGGAGGAAGAGCATTACAGAGTTTCCAGCAGC  
AGCAGGAAAGCTTTGGTTAATTTGGAAATGGATGATAGCATTAAAATAACAGAAGC  
GCCTCCAGGTCTCTGAAGCTTCAGTCCCCCAGCTGAAAGCCAGAAAAGACTAAGCC  
CACTAAGCCTTTTGATCCCTTTGGAAGCAAAGAAGCTTTCCTTCCCTGGGGTGAAGAC  
TCTCCTCAGAAGATTTCTGTCTCTGCCTATGTTACAAGAGGAATCAAAACCAAGAC  
AGAAGAGCTCAGGATGCAGGTGAGAGGCAGGGAAGTCAGCGGCTTGTTGATCTCCC  
TCACTGCTGTCTGCCTGGTGGTCACCCCTGGGAGCAGGGCCTGTCCTCGCCGCTGTG  
CCTGCTATGTGCCACAGAGGTGCACTGTACATTTCCGGTACCTGACCTCCATCCCAG  
ATGGCATCCCGGCCAATGTGGAACGAATAAATTTAGGATATAACAGCCTTACTAGAT  
TGACAGAAAACGACTTTGATGGCCTGAGCAAAGCTGGAGTTACTCATGCTGCACAGT  
AATGGCATTACAGAGTCAGTGACAAGACCTTCTCGGGCTTGCAGTCCTTGCAGGTC  
TTAAAAATGAGCTATAACAAAGTCCAAATCATTTCGGAAGGATACTTTCTACGGACTC  
GGGAGCTTGGTCCGGTTGCACCTGGATCACAACAACATTGAATTCATCAACCCTGAG  
GCCTTTTATGGACTTACCTCGCTCCGCTTGGTACATTTAGAAGGAAACCGGCTCACA  
AAGCTCCATCCAGACACATTTGTCTCATTAAGCTATCTCCAGATATTTAAAACCTCTT  
TCATTAAGTACCTGTTCTTGTCTGATAACTTCCTGACCTCCCTCCCAAAGAAATGGT  
CTCCTACATGCCAAACCTAGAAAGCCTGTATTTGCATGGAAACCCATGGACCTGTGA  
CTGCCATTTAAAGTGGTTGTCTGAGTGGATGCAGGGAAACCCAGATATAATAAAAT  
GCAAGAAAGACAGAAGCTCTTCCAGTCCTCAGCAATGTCCCCTTTGCATGAACCCCA  
GGATCTCTAAAGGCAGACCCTTTGCTATGGTACCATCTGGAGCTTTCCTATGTACAA  
AGCCAACCATTGATCCATCACTGAAGTCAAAGAGCCTGGTTACTCAGGAGGACAAT  
GGATCTGCCTCCACCTCACCTCAAGATTTTCATAGAACCCTTTGGCTCCTTGTCTTTGA  
ACATGACANANNTNTCTGGAAATAAGGCCGACATGGTCTGTAGTATCCAAAAGCCA  
TCAAGGACATCACCAACTGCATTCCTGAAGAAAATGACTACATCATGCTAAATGC  
GTCATTTTCCACAAATCTTGTGTGCAAGTGTAGATTATAATCACATCCAGCCAGTGTG  
GCAACTTCTGGCTTTATACAGTGAAGTCTCCTCTGATACTAGAAAGGAAGCCCCAGCT  
TACCGAGACTCCTTCACTGTCTTCTAGATATAAACAGGTGGCTCTTAGGCCTGAAGA  
CATTTTTACCAGCATAGAGGCTGATGTCAGAGCAGACCCTTTTGGTTCCAACAAGA  
AAAAATTGTCTTGCAGCTGAACAGAACTGCCACCACACTTAGCACATTACAGATCCA  
GTTTTCCACTGATGCTCAAATCGCTTTACCAAGGGCGGAGATGAGAGCGGAGAGAC  
TCAAATGGACCATGATCCTGATGATGAACAATCCCAAAGTGAACGCACTGTCTGG  
TTGGCGGCACTATTGCCCTGAGCTGTCCAGGCAAAGGCGACCCTTACCTCACTTGG  
AATGGCTTCTAGCTGATGGGAGTAAAGTGAGAGCCCCTTACGTTAGCGAGGATGGG

CGAATCCTAATAGACAAAAATGGGAAGTTGGAAGTGCAGATGGCTGACAGCTTTGA  
TGCAGGTCTTTACCACTGCATAAGCACCAATGATGCAGATGCGGATGTTCTCACATA  
CAGGATAACTGTGGTAGAGCCCTATGGAGAAAGCACACATGACAGTGGAGTCCAGC  
ACACAGTGGTTACGGGTGAGACGCTCGACCTTCCATGCCTTTCCACGGGTGTTCCAG  
ATGCTTCTATTAGCTGGATTCTTCCAGGGAACACTGTGTTCTCTCAGCCATCAAGAG  
ACAGGCAAATTCTTAACAATGGGACCTTAAGAATATTACAGGTTACGCCAAAAGAT  
CAAGGTCATTACCAATGTGTGGCTGCCAACCCATCAGGGGCCGACTTTTCCAGTTTT  
AAAGTTTCAGTTCAAAAAGAAAGGCCAAAGGATGGTTGAGCATGACAGGGAGGCAG  
GTGGATCTGGACTTGGAGAACCCAACCTCCAGTGTTTCCCTTAAGCAGCCAGCATCTT  
TGAAACTCTCTGCATCAGCTTTGACAGGGTCAGAGGCTGGAAAACAAGTCTCCGGTG  
TACATAGGAAGAACAACATAGAGACTTAATACATCGGCGGCGTGGGGATTCCACG  
CTCCGGCGATTTCAGGGAGCATAGGAGGCAGCTCCCTCTCTCTGCTCGGAGAATTGAC  
CCGCAACGCTGGGCAGCACTTCTAGAAAAAGCCAAAAAGAATTCTGTGCCAAAAAA  
GCAAGAAAATACCACAGTAAAGGCCAGTGCCACTGGCTGTTCCCTCGTGGAACCTCA  
CTGACGAGGAAAAGGATGCCTCTGGCATGATTCTCCAGATGAAGAATTCATGGTTC  
TGAAAATAAGGCTTCTGGTGTCCCAGGAAGGTCACCAACTGCTGACTCTGGACCAG  
TAAATCATGGTTTTATGACGAGTATAGCTTCTGGCACAGAAGTCTCAACTGTGAATC  
CACAAACACTACAATCTGAGCACCTTCTGATTTCAAATTATTTAGTGTAACAAACG  
GTACAGCTGTGACAAAGAGTATGAACCCATCCATAGCAAGCAAAATAGAAGATACA  
ACCAACCAAAACCAATCATTATCTTTCCATCAGTAGCTGAAATTCGAGATTCTGCT  
CAGGCAGGAAGAGCATCTTCCCAAAGTGACACACCCTGTAACAGGGGGAAACATGGC  
TACCTATGGCCATACCAACACATATAGTAGCTTTACCAGCAAAGCCAGTACAGTCTT  
GCAGCCAATAAATCCAACAGAAAGTTATGGACCTCAGATACCTATTACAGGAGTCA  
GCAGACCTAGCAGTAGTGACATCTTCTCACACTACTGCAGACCCTAGCTTCTCCA  
GTCACCCTTCAGGTTTCACACACCACTGCCTCGTCTTTATTTACATTCTAGAAACAA  
CAATACAGGTAACCTTCCCTTGTCCAGGCACTTGGGAAGAGAGAGGACAATTTGGA  
GCAGAGGGAGAGTTAAAAACCCACATAGAACCCCAAGTTCTCCGACGGCATAGACAC  
AGGACTGTGAGGCCAGCAATCAAGGGACCTGCTAACAAAAATGTGAGCCAAGTTCC  
AGCCACAGAGTACCCTGGGATGTGCCACACATGTCTTCCGCAGAGGGGGCTCACAG  
TGGCTACTGCAGCACTGTCAGTTCCAAGTTCATCCACAGTGCCCTCCCCAAAATA  
ATAATGTTGGGGTCATAGCAGAAGAGTCTACCACTGTGGTCAAGAAACCACTGTTAC  
TATTTAAGGACAAACAAAATGTAGATATTGAGATAATAACAACCACTACAAAATAT  
TCCGGAGGGGAAAGTAACCACGTGATTCTACGGAAGCAAGCATGACTTCTGCTCC  
AACATCTGTATCCCTGGGGAAATCTCTGTAGACAATAGTGGTCACTGAGCATGCC  
TGGGACCATCCAAACTGGGAAAGATTCAGTGGAACAACACCACTTCCAGCCCCC  
TCAGCACACCCTCAATACCAACAAGCACAAAATTCTCAAAGAGGAAAACCTCCCTTG  
CACCAGATCTTTGTAAATAACCAGAAGAAGGAGGGGATGTTAAAGAATCCATATCA  
ATTCGGTTTACAAAAGAACCCAGCCGAAAGCTTCCCAAAATAGCTCCTCTTTTACC  
CACAGGTCAGAGTTCCCCCTCAGATTCTACAACCTCTTTGACAAGTCCGCCACCAGC  
TCTGTCTACAACAATGGCTGCCACTCAGAACAAGGGCACTGAAGTAGTATCAGGTG

CCAGAAGTCTCTCAGCAGGGAAGAAGCAGCCCTTCACCAACTCCTCTCCAGTGCTTC  
CTAGCACCATAAGCAAGAGATCTAATACATTAACCTTCTTGTCAACGGAAACCCCA  
CAGTGACAAGTCCTACTGCTACTGCATCTGTCAATTATGTCTGAAACCCAACGAACAA  
GATCCAAAGAAGCAAAAGACCAAATAAAGGGGCCTCGGAAGAACAGAAACAACGC  
AAACACCACCCCCAGGCAGGTTTCTGGCTATAGTGCATACTCAGCTCTAACAACAGC  
TGATACCCCCTTGGCTTTTCAAGTCATTCCCCACGACAAGATGATGGTGGAAATGTAAG  
TGCAGTTGCTTATCACTCAACAACCTCTCTTCTGGCCATAACTGAACTGTTTGAGAA  
GTACACCCAGACTTTGGGAAATACAACAGCTTTGGAAACAACGTTGTTGAGCAAAT  
CACAGGAGAGTACCACAGTGAAAAGAGCCTCAGACACACCACCACCCTCCTCAGC  
AGTGGGGCGCCCCCAGTGCCCACTCCTTCCCCACCTCCTTTTACTAAGGGTGTGGTT  
ACAGACAGCAAAGTCACATCAGCTTTCCAGATGACGTCAAATAGAGTGGTCACCAT  
ATATGAATCTTCAAGGCACAATACAGATCTGCAGCAACCCTCAGCAGAGGCTAGCC  
CCAATCCTGAGATCATAACTGGAACCACTGACTCTCCCTCTAATCTGTTTCCATCCAC  
TTCTGTGCCAGCACTAAGGGTAGATAAACCACAGAAATTCTAAATGGAAGCCCTCTCC  
CTGGCCAGAACACAAATATCAGCTCAAGTCATACTCCGAAACCATTGAGAAGGGCA  
AAAGGCCAGCAGTAAGCATGTCCCCCACCTCAGCCTTCCAGAGGCCAGCACTCAT  
GCCTCACACTGGAATACACAGAAGCATGCAGAAAAGAGTGTTTTTGATAAGAAACC  
TGGTCAAAACCCAACCTTCCAAACATCTGCCTTACGTCTCTCTACCTAAGACTCTATTG  
AAAAAGCCAAGAATAATTGGAGGAAAGGCTGCAAGCTTTACAGTTCCAGCTAATTC  
AGACGTTTTTCTTCCCTTGTGAGGCTGTTGGAGACCCACTGCCCATCATCCACTGGACC  
AGAGTTTCATCAGGANTTGAAATATCCCAAGGGACACAGAAAAGCCGGTTCCACGT  
GCTTCCCAATGGCACCTTGTCCATCCAGAGGGTCAGTATTCAGGACCGTGGACAGTA  
CCTGTGCTCTGCATTTAATCCACTGGGCGTAGACCATTTTCATGTCTCTTTGTCTGTG  
GTTTTTTACCCGGCAAGGATTTTGGACAGACATGTCAAGGAGATCACAGTTCACCTT  
GGAAGTACTGTGGAATAAAGTGCAGAGTGGAGGGTATGCCGAGGCCTACGGTTTC  
CTGGATACTTGCAAACCAAACGGTGGTCTCAGAAACGGCCAAGGGAAGCAGAAAGG  
TCTGGGTAACACCTGATGGAACATTGATCATCTATAATCTGAGTCTTTATGATCGTG  
GTTTTTACAAGTGTGTGGCCAGCAACCCATCTGGCCAGGATTCAGTGTGGTTAAGA  
TACAAGTCATCACAGCTCCCCCTGTCATTATAGAGCAAAAGAGGGCAAGCCATCGTTG  
GGGTTTTAGGTGGAAGTTTGAAACTGCCCTGCACTGCAAAAGGAACTCCCCAGCCTA  
GTGTTCACTGGGTCCTTTATGATGGGACTGAACTAAAACCATTGCAGTTGACTCATT  
CCAGATTTTTCTTGTATCCAAATGGAACCTCTGTATATAAGAAGCATCGCTCCTTCAGT  
GAGGGGCACTTATGAGTGCATTGCCACCAGCTCCTCAGGCTCAGAGAGAAGGGTAG  
TGATTCTTACTGTGGAAGAGGGAGAGACAATCCCCAGGATAGAACTGCCTCTCAG  
AAATGGACTGAGGTGAATTTGGGTGAGAAATTACTACTGAACTGCTCAGCTACTGG  
GGATCCAAAGCCTAGAATAATCTGGAGGCTGCCATCCAAGGCTGTCATCGACCAGT  
GGCACAGAATGGGCAGCCGAATCCACGTCTACCCAAATGGATCCTTGGTGGTTGGG  
TCAGTGACGGAAAAAGACGCTGGTGACTACTTATGTGTGGCAAGAAACAAAATGGG  
AGATGACCTAGTCCTGATGCATGTCCGCCTGAGATTGACACCTGCCAAAATTGAACA  
GAAGCAGTATTTTAAGAAGCAAGTGCTCCATGGGAAAGATTTCCAAGTTGACTGCA

11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100

AGGCCTCTGGCTCCCCTGTGCCTGAGGTATCCTGGAGTTTGCCTGATGGGACAGTGC  
TCAACAATGTAGCCCAAGCTGATGACAGTGGCTATAGGACCAAGAGGTACACCCTT  
TTCCACAATGGAACCTTGTATTTCAACAACGTTGGGATGGCAGAGGAAGGAGATTAT  
ATCTGCTCTGCCCAGAACACCTTAGGGAAAGATGAGATGAAAGTCCACCTAACAGT  
TCTAACAGCCATCCCACGGATAAGGCAAAGCTACAAGACCACCATGAGGCTCAGGG  
CTGGAGAAACAGCTGTCCTTGACTGCGAGGTCACTGGGGAACCGAAGCCCAATGTA  
TTTTGGTTGCTGCCTTCCAACAATGTCATTTCACTTCTCCAATGACAGGTTACATTTT  
ATGCCAATAGAACTTTGTCCATCCATAAAGTGAAACCACTTGACTCTGGGGACTATG  
TGTGCGTAGCTCAGAATCCTAGTGGGGATGACACTAAGACATACAACTGGACATT  
GTCTCTAAACCTCCATTAATCAATGGCCTGTATGCAAACAAGACTGTTATTAAAGCC  
ACAGCCATTTCGGCACTCCAAAAAATACTTTGACTGCAGAGCAGATGGGATCCCATCT  
TCCCAGGTACAGTGGATTATGCCAGGCAATATTTTCTCCAGCTCCATACTTTGGA  
AGCAGAGTCAACGGTCCATCCAAATGGAACCTTGGAGATGAGGAACATCCGGCTTTC  
TGACTCTGCGGACTTCACCTGTGTGGTTTCGGAGCGAGGGAGGAGAGAGTGTGTTGGT  
AGTGCAGTTAGAAGTCCTAGAAATGCTGAGAAGACCAACATTCAGAAACCCATTCA  
ACGAAAAAGTCATCGCCCAAGCTGGCAAGCCCGTAGCACTGAACTGCTCTGTGGAT  
GGGAACCCCCACCTGAAATTACCTGGATCTTACCTGACGGCACACAGTTTGCTAAC  
AGACCACACAATTCCCCGTATCTGATGGCAGGCAATGGCTCTCTCATCCTTTACAAA  
GCAACTCGGAACAAGTCAGGGAAGTATCGCTGTGCAGCCAGGAATAAGGTTGGCTA  
CATCGAGAACTCATCCTGTTAGAGATTGGGCAGAAGCCAGTCATTCTGACATACGA  
ACCAGGGATGGTGAAGAGCGTCAGTGGGGAACCGTTATCACTGCATTGTGTGTCTG  
ATGGGATCCCCAAGCCAAATGTCAAGTGGACTACACCGGGTGGCCATGTAATCGAC  
AGGCCTCAAGTGGATGGAAAATACATACTGCATGAAAATGGCACGCTGGTCATCAA  
AGCAACAACAGCTCACGACCAAGGAAATTATATCTGTAGGGCTCAAAACAGTGTTG  
GCCAGGCAGTTATTAGCGTGTGAGTGTGTTGTGGCCTACCCTCCCCGAATCATAA  
ACTACCTACCCAGGAACATGCTCAGGAGGACAGGGGAAGCCATGCAGCTCCACTGT  
GTGGCCTTGGGAATCCCCAAGCCAAAAGTCACCTGGGAGACGCCAAGACACTCCCT  
GCTCTCAAAAGCAACAGCAAGAAAACCCCATAGAAGTGAGATGCTTCACCCACAAG  
GTACGCTGGTCATTTCAGAATCTCCAAACCTCGGATTCCGGAGTCTATAAGTGCAGAG  
CTCAGAACCTACTTGGGACTGATTACGCAACAACCTTACATCCAGGTACTCTGACAGG  
AAGGGGGAGACTAAAATTCAACAGAAAGTCCACATCCACAGGGTTTATTTTTTGGA  
GAAGTTTAATCAAAGGCAGCCATAGGCATGTAAATGAGTCTGAATACATTTACAGT  
ATTAAATTTACAATGGACATGCGATGAGACTTGTAATGAAAGCATTGTGAACTGA  
AACCGAGTCTCTGTGGATCTCAAAGCAAACCTCTTAACCTTAAGGCACTTTGATTTTGC  
CAACAAATAATAACAAACATTAAGAGAAAAAAATGATCCACTACGAAATAACAAAC  
GGCTAATGCACCTGAATTCTCAGTAAAAAGACCTTTCTCTCGCTAACAGTTGCCAGC  
TGCTTCGTGTCTGTTTCTACCAATGTCAACAAACATCGCACACAGGGTGAATGGAGT  
CAACGGGAAAGATTAAAGTTTGCAGTCTGTGTAAATCTCAATGTACAAATATTCTGTC  
NCTGGTTTATAAACATTTTGATAAAACCGAAAAA  
AAAA (SEQ ID NO: 1)

Figure 2

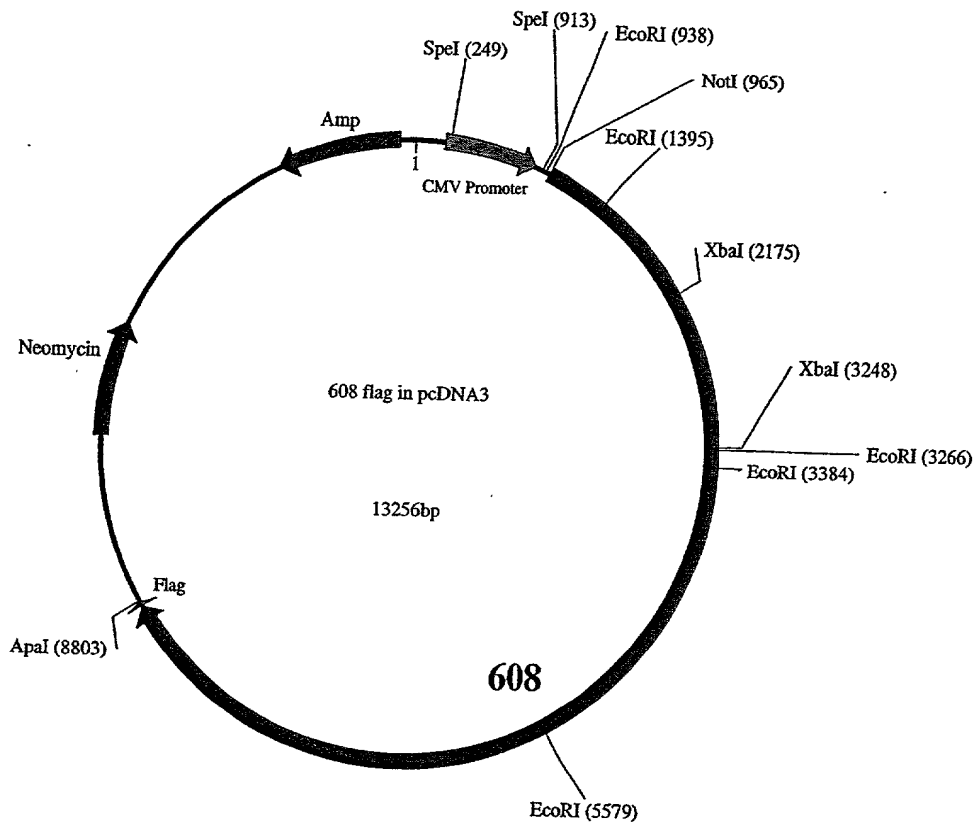


Figure 3

MQVRGREVSGLLISLTA VCLVVT PGSRACPRRCACYVPTEVHCTFRYLTSIPDGIPANVE  
RINLGYNLTRLTENDFDGLSKLELLMLHSNGIHRVSDKTFSGLQSLQVLKMSYNKVQII  
RKDTFYGLGSLVRLHLDHNNIEFINPEAFYGLTSLRLVHLEGNRLTKLHPDTFVSLSYLQI  
FKTSFIKYLFLSDNFLTSLPKEMVSYMPNLESYLHGNPWTCDCHLKWLSEWMQGNPDI  
IKCKKDRSSSSPQQCPLCMNPRISKGRPFAMVPSGAFLCTKPTIDPSLKSLSVTQEDNGS  
ASTSPQDFIEPFGSLSLNMTXXSGNKADMVCSIQKPSRTSPTAFTEENDYIMLNASFSTNL  
VCSVDYNHIQPVWQLLALYSDSPLILERKPQLTETPSLSSRYKQVALRPEDIFTSIEADVR  
ADPFWFQOEKIVLQLNRTATTLSTLQIQFSTDAQIALPRAEMRAERLKWTMILMMNNPK  
LERTVLVGGTIALSCPGKGDPSPHLEWLLADGSKVRAPYVSEDGRILIDKNGKLELQMA  
DSFDAGLYHCISTNDADADVLT YRITVVEPYGESTHDSGVQHTVVTGETLDLPCLTSTGV  
PDASISWILPGNTVFSQPSRDRQILNNGTLRILQVTPKDQGHYQCVAANPSGADFSSFKV  
SVQKKGQRMVEHDREAGGSGLGEPNSSVSLKQPASLKSASALTGSEAGKQVSGVHRK  
NKHRLIHRRRGDSTLRRFREHRRQLPLSARRIDPQRWAALLEKAKKNSVPKKQENTTV  
KPVPLAVPLVELTDEEKDASGMIPPDEEFMVLKTKASGVPGRSPTADSGPVNHGFMSTI  
ASGTEVSTVNPQTLQSEHLPDFKLFSVTNGTAVTKSMNPSIASKIEDTTNQNPIIFPSVAE  
IRDSAQAGRASSQSAHPVTGGNMATYGHNTYSSFTSKASTVLQPINPTESYGPQIPITGV  
SRPSSDISSHTTADPSFSSHPSGSHTTASSLFHIPRNNNTGNFPLSRHLGRERTIWSRGRV  
KNPHRTPVLRHRHRHRTVRPAIKGPANKNVSQVPATEYPGMCHTCPSAEGTLVATAALS  
VPSSSHSALPKTNNGVIAEESTTVVKKPLLLFKDKQNV DIEITTTTKYSGGESNHVIPTE  
ASMTSAPTSVSLGKSPVDNSGHLSPGTIQTGKDSVETTPPLSPPLSTPSIPTSTKFSKRKTP  
LHQIFVNNQKKEGMLKNPYQFGLQKNPAAKLPKIAPLLPTGQSSPSDSTLLTSPPPALST  
TMAATQNKGT EVVSGARSL SAGKKQPFTNSSPVLPTISKRSNTLNFLSTETPTVTSPTAT  
ASVIMSETQRTRSKEAKDQIKGPRKNRNNANTTPRQVSGYSAYSALTADTPLAFSHSP  
RQDDGGNVSAVAYHSTTSLLATTELF EKYTQTLGNTTALETTLLSKSQESTTVKRASDTP  
PPLLSSGAPPVPTSPPPFTKGVVTD SKVTSAFQMTSNRVVTIYESSRHNTDLQQPSAEAS  
PNPEIITGTTDSPSNLFPSTSVPALRVDKPQNSKWKPSWP EHKYQLKSYSETIEKGKRPA  
VSMSPHLSLPEASTHASHWNTQKHA EKS VFDKKPGQNPTSKHLPYVSLPKTLLKKPRIIG  
GKAASFTVPANS DVFLPCEA VGDPLPIIHWTRVSSGXEISQGTQKSRFHVLPNGTLSIQRV  
SIQDRGQYLCSAFNPLGVDFH FVSLSVVFYPARILDRHVKEITVHFGSTVELKCRVEGMP  
RPTVSWILANQTVVSETAKGSRKVWVTPDGTLIYNLSLYDRGFYKCVASNPSGQD SLL  
VKIQVITAPPVIEQKRQAI VGVLGGS LKLPCTAKGTPQPSVHWVLYDGT ELKPLQLTHS  
RFFLYPNGTLYIRSIAPSVRGT YECIATSSSGSERRVILTVEEGETIPRIETASQKWTEVN  
LGEKLLLNC SATGDPKPRIIWR LPSKAVIDQWHRMGSRIHVYPNGSLVVGSVTEKDAGD  
YLCVARNKMGGDLVLMHVR LRLTPAKIEQKQYFKKQVLHGKDFQVDCKASGSPVEV  
SWSLPDGTVLNNVAQADDSGYRTKRYTLFHNGTLYFN NVGMAEEGDYICSAQNTLGK  
DEMKVHLTVLTAIPRI RQSYKTTMRLRAGETA VLDCEVTGEPKPNVFWLLPSNNVISFS  
NDRFTFHANRTL SIHKVKPLDSGDYVCVAQNPSGDDTKTYKLDIVSKPPLINGLYANKT  
VIKATAIRHSKKYFDCRADGIPSSQVTWIMPGNIFLPAPYFGSRVTVHPNGTLEMRNIRLS

1000 900 800 700 600 500 400 300 200 100 0

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

Figure 4

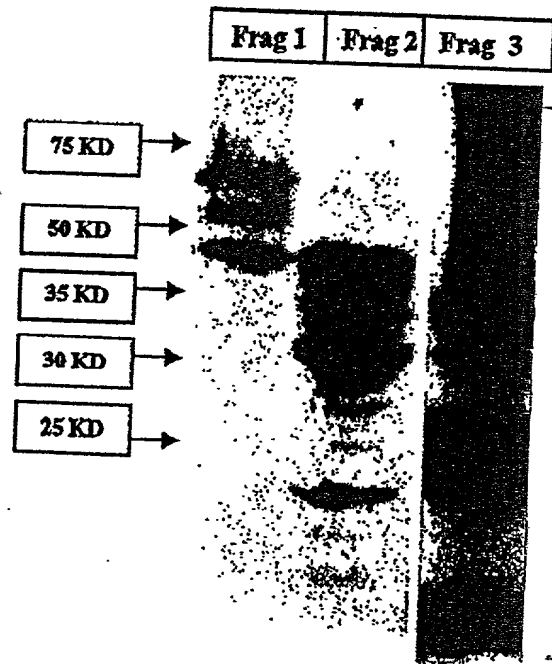




Figure 7

Nuc 1

↓  
TTTGGAAACCAACCCAGATGCCCTCAACAGAGAAATGGGCCAGAAAATGTGGTCCATTTA  
TCCAATGGAAATACTACTCAACTTATTAATAAACCAACGACTTTCATAAAATTTTAGGCAAAAT  
GNATGGTCTGNAGGATCTTGAGTGAGGTAAACCAATCACAAAAGAACACTCATGGTATGC  
ACTCACTGATAAGTGGCTATTGTCTATGGAGTGATTAAAGGGAAGAAGACACATAGC  
TTTTTGTGTGTATAATATTAAGATGGAAATTTGCCAGTGTCTTTGGCTTATGAGTGAATCT  
TGTTTCAGTGGATTACCGGAAGAAAATAATAAGTGAAGTGTAGGAAGAAGTAGTTAATCA

↓ Nuc 390

AGGTGACAAAGTATCCTGACACATTGGGAAAAGACCACAGTCCAGGAAACTGAGTCTTAA  
GGATTTCATATTAACCTCCAGTTCCTTACAGAGTCTGAGACTTTGGCAGATCAGACACTT  
AACTTCAACAGCTTCCTACACAGAGCAGTTACTATCCTTGCACCTTCACACATGGAGTGTGA  
CCATTAAAGCTGCACCTGAAACATGAGTCTGACTTGTTAATAATCTTAAATAACAAATTGTGT  
TGTAAGTATGTGACCAAGAGCATGGTTCATGCTATTAACCTTTGATGTTCTATGGACTCT  
TAATTTTATGGTAGAAATGTCAACAAGCTTGTGGAGGCTGGAAAGATACAAGGCTTAAGAG  
GATGGCTTTTCACTTTTGAAGTAAATTCAGTATGTGTTCTGGCATCCCTTTTCTTAAAGCA  
ATTAAACCCCCCAAGTAGGCATAATTTTAATGCTTACTTCATCAGAATATGTCTAATTGAC  
TCTTCTAAAAAGACTTTGGTATGCATAGGATCTAAATGTAAATGTGATTTACTGACATAAT  
AAATAGGAGAAACTGAGCTAGAATAGGTATAAAATATGTGCTGGCTTTCTAATAGGTCTT  
ATAGGTATATATAAGAGGTGGGAAAGGAATATTTGAAACATCTAGAAGTAAATGATCCTG  
AGTAGGATCCTGGGAAATACGTACTCTAACACACTGCAATCATCTCTGTGGTTTGTCT  
GGAGCTGAGGTCTGGAAAGGCTCGACCTTGGTTAGAAATAACCTACCGAATACAGAGCTAT  
GACGTTAGTCTGGAAGGAGCTTTGGAAGAATGACAAGCTGTAGCTGCCAGAACATATA  
GATGCCATATTTCCAAGGCAAGTGTCCACATGCCGACATCTTAAGAATATGGTTGTCTCTG  
CAGTGTCTAAGGACCTTGTCTGTCACACAGGTCTCCAGGGTTAGTGTCTAAGTCTGACTGC  
TTGACTCTTTAATCTAOCCTTGATCATTATGACTAGAAATCACTTGGTGATTAGCAACTG  
GATATGGAATATTACTAATTTGTACCCAAAGCCAGGCCACCTCAGCTTTGGCAGCTTCATTC  
ATTCTGTGGAGCCCAAGTCAAGTGGGTTTGAATCAAACTGTACTGTTTCTACTTACAAGACGC  
ATTACCTGAGATGAGTCAATTTTCTTCAAGTCTTTTGAAGAGTCAATTAGACATATT  
CTGATGAAGTAAGCATATAAAGTGAGAGCAGCATGAATGTGTTCCATGTATGCTCATGGA  
TGCTATTATAATGTGGAATAAACTGACTTAAAAAAGGCTTATGATACCTTGTCTACA  
GAGTAAATCTTCCATAAATATCATCTGCATTTATAAATTTATTTTATAATCCATCAATTA  
AAACCTTTAGAAATTTTGTAAACACAAAGATCCCTAGGCCCTGCCCTAGGATGGTCTGTA  
TGCTGGGCTGAGAGATGGAGCTTAAGAACTTACTTGTCTCCAGGAGCACATCTTCAGAAC  
ATCTGCCCTCAAAACATTTATCCCAAATGCTCATCAAAAGGCTCACTCACATGTGCTTCAACC  
ACAGGGATTAACAGTCAATTTTGTACATTTTCTCAAAAGGCTGGAAAGCCTGCTAGAGGAA  
CAGGATGTATCAGGATAACATCCAACTTACAAAAGGATGTATTAACCTCACCAACAACA  
AACAAACAAGCAACAACCCATAAAATTTATCAAGGCAAAATGAACCTAAGCCATATGCA  
GAAAAAGTATTATATGTTCTCATTTGTGGGTGTTTTTCTTAAATAGTCAAAATATGAGAAT  
ATAGACAAAAGATGGTTTATGCAAGTGGGGATGGCGAAGGATACTTGTAGATTAGAGGACA  
CAAAGCAACAACCTACAGAGTGAAGTAATCCAGAGACTTAATGTATAATATGAGGACTGTA  
TTAATAATTTCTATTTAAGATACACAGCAAAAGAGTGTATCTTACTAACACACACACTTAC  
ATAGAGAGAATAAAGTGATAGATACGTTTGTGTTTATCTTCTATGTAGCTGATAATTTATAT  
TGTACACCTCAAAACATAGATAACCAACAAGAGGAAGAGGATAGGTGCTCTCCAGGGC  
GGAAGAGTACATTGAAAGTACAGACACCATTTGTGTAGATGTACCACATGGAGGAGCTAGA  
GAAAGTAGCCAAAGGAGCTAAAGGGATCTGCAACCTATAGGTGGAACAACATTTATGAGCT  
AACCAGTACCCCGAGCTCTTGAAGTCTAGCTGCATATATCAAAAAGATGGCCTAATCGG  
CCATCACTGGAAAGAGAGGCCATTTGACTTGCACAACTTATATGCCCCAGTACAGGGGA  
ATACCAGGGCCAAAAGGGGGAGTGGGTGGCCAGGGGAGTGGGGGTGGGTGGATATGGG  
GGACTTTTGGTATAGCATTGGAAATGTAAATGAGTTAAATACCTAATAAAAAATGGAAAA

↑ Nuc 2817

AAAAAAAAAAAAAAAAAAAAAAAAAGGAAGGTGAGACACCTCACTTCACTGCTATCTCAACTT  
GCAACAGAAAGGGAGTCACAAACCCAGGACAAACACAGTGATTGAAGCGTCTTTGAA  
TGTTATTGCTGTTGTTGTTTACCACCATCTAAGCATATATTCATTGTGAAAACTACGGGT  
CTATGACATGTTTTTTTATTCAAGTATATCACATGCTGTACGCATATTTGGCACCCTACCA

[illegible]

▼ Exon 1 (Nuc 3661)

▲ Nuc 4203

TGTAAATACAGTTTGAGGCTTAAGTGTACGGGAACCTCATGTGGTATTCAITTAACGGCTCT  
 CTCTCTTATAACTAAGCTCTTAAGGTGCATATAGTCTCTCTGTTTCCAGCTACCTTGTAC  
 CATCTTTGTTTATCTAATAATAGCAAGCTCATCTGCTTTTAAATCATCAGCGAGAGAGTATT  
 CAAAAATATTCAGTGTATTAACAGTGCAGTGTGGCATAGAAGTAATCAATTAGTAAATC  
 TAAATTTGGGTTAAACATCTTTCATTAACGAGTCCAGGTTGGGAGGGATCACTGAGGCTTCGC  
 CACGTGCGGGTTAAAGATATTTCTAACAAGAGAAGCAGAATTCCTCTCTTGGGCATGCTCC  
 CCATCACGTGTGCAGTAAGCAGAGGGGTGTTTCCAAGCAGAGAAAGAGCAGACAGTGTTA  
 TGCTTCCAAAGTCAGAGAGCTACAGCCCTCCAGCTGGTCAGTTTACTGTCTCTCCGGTCAIT  
 AGTTGGCTCTGAAAAGGCCCATGTGTCTTTATTGGCAAGGACTTGCAGACATGCTAGAAA  
 GAAATTTGAOCTTTTTCCTAGTGGGTTATTACAGCTGTAAAAGTATTTGGAAAGGTTAAG  
 CCAAAATAAAATAAAACACATATTTAAATAATACAATGTTACAAAAATGTATCATATAAAGAA  
 GTACATTCATAAATGCAATGTGAAAAATATATATAATTTTATCTATTTACTGGTGCAAAG  
 TTTTCTAAATGTCACATGTACTATTTTATATTTATAAAAAATATTTTAAAAATGTATATAAA  
 AGTGTAAAGGCTCTTGGTCAACACAGAGAGTAAATTTTACAAAGCTTTAATTTGTCCTCCAT  
 AAGCTATTATTATGATCTCTAATGACAGGGAGCTCTGTCTTTTATTTGGGAAATGAGAAGCTATG  
 AAGATATGTTTACAATAATAAAGCCATTTAGTGATAAAGTCCAATTGGGAAGCTAGCACAC  
 ACTGGTTTATAAAGAGAACAGTTCCTGTAGTCTATGCAAGTTTACACTCTAGGGAATAAG  
 AGTCTCTCTTCTCCAGATTTCACTAGCATTTGTTGTATCATTTATCTCTCTGTATGATGAG  
 CATTTACAGTGGGAATAAGATTAGGATGTCTCAAAGGAATGTCAATTTGGATGCCCTGAACAA  
 CTCTTCAAGGCTCTTCTCTCAGTCTACTAGTCTATTCATTTATTGGATAATTGGGGGATGGTGT  
 TAATTTTTTTCAGTCTCTATGGAATTTCCAAAAACAAAAACAAAAACAAAAACAAAAA  
 AACTCTGAAACTAGAACTACCAATCCATTACTGGGTATGTAACAAAAGAGAAATCTGCACA  
 GAATTTATTGCTACATTTGTTCAATTATTTCAAGACAGCCAGAATGTGGAAACCAACTTACGTA  
 GCGGTCAAAAATATGAACGGGATAAAGAAAAATGTGGAAATGTGTACAACAGAGTCCCATGTG  
 GCCATAAAAGAGTGAAGATCATGACATATGCAGGAATCCGATGCACTGGAAATCAAAATG  
 GCGTAATCAAAAACGAAGCAGACTCAAAAGGAAACACCGGTGACCTTCTGTACAACACA  
 GAAGCTAGATTACACTGTGACGTGGCGCATGTGTGTTTGAATTTATTTAGTATACACT  
 ATTCTAATCTGTGAGTGTGTATAAAGGCATGCATGTAAAGCAAAAAACAAGCTAGCTGGG  
 TGGTATGAGAGAGAAAGCAATGAGAGGAGTTAAATAAGAACGAAGCATAGTAACATAGGTG  
 CCAGGATGAAATGCATTAATTTGTATGCTAACTAAACCCAGACAGAGGAGGCACAGCTTCA  
 AACCAGGTTGAAATCCAGCAGCAGAGAGAGGGGAAGTAGACACAAAGTTTCGCCACTAAC  
 CAAGAAGCCATTTGCAGTTGCTGCTGCTGGGAGGGGCGTCCAGTTTCTCCAGTCTGAC  
 ACTGTGTATAACAACAGTGTGACAATACAAAGTTGGCATGATGGATGGTTTTTGTGCTATT  
 TTTCATTTTTCCTACTGTTTGTGTTGTTGGTGTGTTGTTGGTGGTGGCTGTGGTTTTT  
 ATTGTTTCTTTGAGAGGAGAGAAGGAACATGAAATTTGGTGGGTGGGATGGAAGCTGGAAACG  
 ATCTGGAAGAAGTTGGGGAAGAGAAAAATTGTATGGAGCATATTTAAACAAACAAACA  
 AACAAAAAAGGTTTCATTTGCCACAAAAAGGTGTGAATTAATTAACAGTTACGACT  
 CTTAAAGAAAAATATTOCCAATTATTCACAGAGTTGCTATGTATGCTGTGCCATAGGACTTTG

▼ Exon 2 (Nuc 6559)

CTTGAACCTGGCCCTATAACCTCTGGTGTGGTGTCTTTTCAGGATGCAGAAGAGAGGCGAGGGAA  
GTCAGCTGCTTGTCTGATCTCCCTCACTGCCATCTGCTGGTGGTCACCCCCTGGGAGCAGGGTC  
TGTOCTCGCGATGTGCTGTCTATGTGOCACAGAGGTGCACTGTACATTTGGGGACCTGAOC  
TCCATCCCAGACGGGCATCCCAGCCAATGTGGAACGAGTCAATTTAGGGTGTGT

↑ Nuc 6755

GGACCTTGCCTGATCTCCTTCTCAGAGAGGGACCACTGATTTTCTGTTACTTTGCCCCC  
AAACACCTGTGATTACTTTTAATAGTTTTCTTCTAAAATGGGTTTCATACAAACCTTATATTG  
TGGAGACAATGAACATTTATCCCAATAGTCTTTTACTAGAAGTTGAAGCCCCCTTTAGTT  
GTTTGGGAGCCTCATAAATATGGGGCAGCTTTATTCTGAATGAATTTAAATGAAAAAGAT  
ACAGTTTCTGTTAAACAATCATTATGATACCAAGGAAGAGGAATTGTCAATTGAATATTTAA  
AAAAGCATTTCTTTTGAATTTATAAATACCCATTACAAAATGGCTTACTTAAAATACTTG  
CCTTACTAAATCTGACAAATTTATGGTGATATTTTGAAGGTTTATGAAAATTTGTTTATGTGT  
ATAAATGCACAAGAAATGGGATATGCCATCACCTATGTGCCATTAGTGAGCATGTACAGT  
ATGCCAAACACTATTGTTCACGTTTGGAGGAAGTAATGGGGGTGGGGGAGCAACAAGGGT  
TATAACCGTATACCCAGTGCCTTGGAAAGCGATTGCAAACAGTAAAGACTGACATTTGTGT  
CTCCCTATGAGGGAGGGGCTTGGGCTGAGCACTTTGCAATGAGCATTGTCTCATTGTGCT  
GGCAGGTTTATGATAACTTGACCCAAAGCTAGAGTCACTGGAGAGGAAGGAACCTTCAACT  
GAGAACATGCCTGAAGAAGATCAGATTATAGGCAGGCTGTGGGGCATTTCCTTAATTAG  
TGATTCATGGGGCAGGGCCAGTCCATTGTTCTGTTGTAACATTTCAGGCACATTTAAAA  
AAAAAACAAGCTGAGCAAGTGTCAAGGAGCAAGTCACTGAGCAGCAGCCCTAATGA  
TCTCTGCATCAGCTCTGCTCCAGGTTCTACCTATTTGAGTTCTGTCTAGCTCCCTA  
CAGTGATGAACAATGATGTGGAAGTATAAGCCAAATAAATCCTTTCTTCCCCAAGCTGTCTG  
TTGGTCATGATGTTTATCAGTGATAATAGTCTCATGAAGATGCTGTGTTTATAACA  
CCTTTGGACATAAATCTGTATCTATAGCTGAGGAAAATGGAGCATAGAAAGTCTCCAGA  
CTACACCAGAGTGTAACTGGGCTGAGCTTAGAATCACACCCAGTGCCTCCACTGCC  
GGGGCTTCTTAACCGGAACACAGTTGTAAAAGGGAATTTCTGTTTGTTCATTTTGACA

Exon 3 (Nuc 8089) ▼

TGTGGACTTTAATTGACGATTTCATCTGAAGCTGAAAATGATTTTTTTCCAGGTATAACAGC  
CTCACTAGATTGACAGAAATGACTTTTCTGGCCTGAGCAGACTGGAGTTACTCATGTGCACA  
GCAATGGCATTACAGAGTCACTGACAAGACCTTCTCGGGCTTGCAGTCTTTCAG

Nuc 8218 ↑

GTGAGATAGGTAGAGGGTGATGGAGGCTGAGAAGAGAGGTGCAACTGTGGGTATATACCC  
AAAAGCTGCTGATTCCCGTGGGAGACATCTATAAGCATTCTATAAACTAGAGGCAGATA  
TCAAGGAAGGATTTCAATTGTAATGCAATTTTATGAGAAAATTTGAATATTAAGAAAATG  
CTGGGGAATAATGCTTACACAATTGCGAGGACCTAATTTAGGATCTCCAATAGCCACATAA  
AAAGCACAGCATGGGGCAGACACCTGCAATTCCTGTCCCTGGAAGCACTGTTCCAGAAAT  
CCCAGAGACTCATTGGCCAAACACTCTATTCAATCAATGAAGTGCATATTCAGTGACAAA  
ACTTGAGCTCAGAACTAATGTGGAAAGCATCAGGAAGACAGCCAACATCTGTCTTACT  
CATGCATGAATAAGGGATCCAGAGAGAAAGGAAGAAAAGGAAGGAAGGAAGGAAGG  
AAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGG  
AAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGG  
GGAGAAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGG  
TTTTATCCATAAAAGGTCATTTCCACCTGTTTGTCTGGAAGTAGAGTGGGATCCCTTATAT  
AAGGGCAGTCTTTAACAATAGTAGCATTTTATAAACCATTACAAATTTTGAAGTTTCTCTAC  
TTTTTATCCTCTACCATCTTCAAACCTGAACTACAATTATTCACAAATGAAGAAAATGC  
TGTAAGAGTTTTTACACACCGAAGTGGGAACTTAAGGATTAGACAAGTCTAACAATGAG  
AATGGGGAGAACAAAAAGAGACTGCACAGGGAGGCCCTTCTCTGCTTATAATCTTGACAC  
TTGAGAAGCTAATTGAAGCTGCATGACTACTCAACTCTTTAAGCAACAATGCTGTTGTTC  
ATGAAAAGCAACAATAAAGTACATATGTCCATAATATTCACTAAAAATTTGCAATGCAGCAC  
ATAATAGCAATCAAAGCAATAACACCCACTGTTACAGAGACTTTAAACATGAAACTGGA  
ACTATGTCTAGTGTTTTGAATAGGGTACATAGTATGCTGTGTCTGTATGTACCAATGTTG  
ATTTAGGTCATCAGACAGCATTTTGAACATGTATCTTCAGGAGGAATCATTCTGTATCT  
GCATGAAATTTCTCACTATGTTTATTCTCTTAGCCAGGTTTTTCTCTGATGGAGAAACATT  
GGGTTTGAAGTTTTTACTCCAGGTAAACATTTAGGGAAAAGCTGTCTATGTTCTCAGTTTGG  
CTTTTATTTATGAGGGATGTTGGTATTCCAGAAAATTTCTCTTTTGAAGAGATTACAATTTA  
GGTCAAAACAGAAAAATATGTAAGAAAGTTATTGTTTTTATAGTATTTATGTTCTTTCTT  
TTTAAAAATGGTATGCTTAGAATAATTAAGATTAGATTAGATTAGATTAGAAAAATAATC  
AGAGAGGGAATTGATGAATGCTAAAGCATCATGAAAAATTTCAAAATTTTTGCTTCTAATT  
CAGAATCAATTAATTCATATTACTATAAAGACAGCAGCCAGATGTGTGCCAGCTGAG  
GAGTGGATAAACTGTGTAACGTGAGTGCTATGTAGAAACAGAAAGGAGTGAAGGGTTGA

TGTGCGCTGCAACATCTTGAAAAACATTCGGCTACATGATGGAAGCCAGGCACAAAAAGCC  
ACATATTGCATGGTTATGTTTATATGAAATGTTTAAAAATACATGGATTCTTAGCAAAACAGA  
GTAAGATGTTACTTAGGGTCAGGAAAAGATTAAAAAAAAAAAAACTATTGATGTGGAATG  
ATCTTAATTTGGGGAAAAGACAATTTCTTAAGACGAAATAGTTGAGGTAGATATAGTTAT  
ATCCCTGTGGATATTGTAATAAAACAGCATGCTGTCTCTGAGAAGGGCTAATGAAGGG  
GCAGGAGGAAAGTGAAATGAGATGGTAGAAAGGAAAGTCATATACCATGGCTTCTCTCGTG  
GGTGAATCTAGATATGTTAATATATTGACATAAAGGAAGGAATTGTTTAGGGAAGGATC  
AAAACCAACAGGAGTGAGGGAGACAATAGGAACCAATGAGAGGCCAAAGTTCATGGTCAA  
TGTGTGTGGAGACACCATAATAAAACTCCTTTTTTGTGTTGCTAACTAAAACCACTAAAATC  
TAAAAACAAAACATTTTTGCACAAGAATTATTTATTATTCAATAAAGATGTTTAAATGGGG  
GAAGTTGAAGTTCATTGATAGTCTCATAAAATCTTAAATGTATTAAACTGCTTTTACGTTT  
TTTATTATTAATTAATCTCTGTCTGTCATTATTATCATCATCATTATCGTCATCATCATCTA  
ATGCTTTTACCATACACAAATGTAGGCAGAAGAGTGTAATCCACTTAGTGAGGCAATCTT  
GGAGAGGGAAAAGGAAGCGGATGCGGGGACAGGGCACACAGGAGGACAGTGAGAGGGAA  
ATGAACAAGAAAAAATGTGGACACATGCACAAAAATTCATAGTCCACTACATTACTTTG  
TATTCTAATAATTAAGAAAAATAATAAACCCATTCTGTGCACTTATCACCCAGGCTCAACAG  
TTATCTTGGCCACAGATCCTGTCTCACTGCATCCTGTCCACCTGAGTCCACTTAGCGTCTG  
AATCCAATCCAGGGCATGATGCTTACTCCTACACAGAACTAAAGATTAAAGAGAGTTTAA  
AAGTAACCATGACATCTCTCTGTTTCCCTTAGCGATAAGTTCCTAATATTATGGCTGCTGT  
GTATGTTCTAATTTCTCTAATATTGTACATTTAGTTGGCAACTACTTTGTTTGAATTGAGT  
TGGAGTTAAGGTCCCATAGGATTAATCTCAACATATTTCTATATTTATAAACTTTTCTCTCT

Exon 4 ? (Nuc 11286) ↓

TTGTGAAAGTTCCTTTGAGAAAAACAAATATGCCATATCTTTCTTTACAGGCTCTAAAAATG  
AGCTATAACAAAGTCCAAATAATTGAGAAGGATACTTTGTATGGACTCAGGAGCTTGACCCGG  
TTGCACCTGGATCACAACAACATTGAGTTTATCAACCCCGAGGCGTTTACGGACTCACCTTGC  
TCCGCTTGGTACATCTAGAAGGAAAACCGGCTGACAAAGCTCCATCCAGACACATTGTCTCTTT  
GAGCTATCTCCAGATATTAAAACTCCTTCAATAGNACCTGTACTTGTATGATAACTTCATTG  
ACCTCCCTCCAAAAAGAAATGGTCTCCTCTATGCCAAACCTAGAAAAGCCTTTACTTGCAATGGAA  
AOCATGGACCTGTGACTGOCATTTAAAGTGTTGTCCGAGTGGATGCAGGGAACCCAGGTA

Nuc 11680 ↑

ACTATCTTGTGTTGTTGTTCTTTTTTATARKAOGTATTTTCTCAATTTCAATTAGAAATGA  
TATCCCAAAAGTCCCCATAAOCCTCCCCCCTTCCCTACCTACCCATTCCCATTTTTTGG  
CCCTGGCATTCCCTGTACTGGGGCATATAAAGTTTGGCGTGCCAATGGACCTCTCTTCC  
AGTGATGGCCAACCTAGGCCATCTTTTGATACATATGCAGCTAGAGTCAAGAGCTCTGGGG  
TACTGGTTAGTTCATAATGTTGTTGCACCTACAGGGTTGAA

Nuc 11967 ↑

(SEQ ID NO:3)

Figure 8

Exon/Intron No.	Exon start	Exon end	Exon length	Intron length
1	1	542	542	2356
2	2898	3094	197	1334
3	4428	4557	130	3068
4	7625	8019	394	

Figure 9

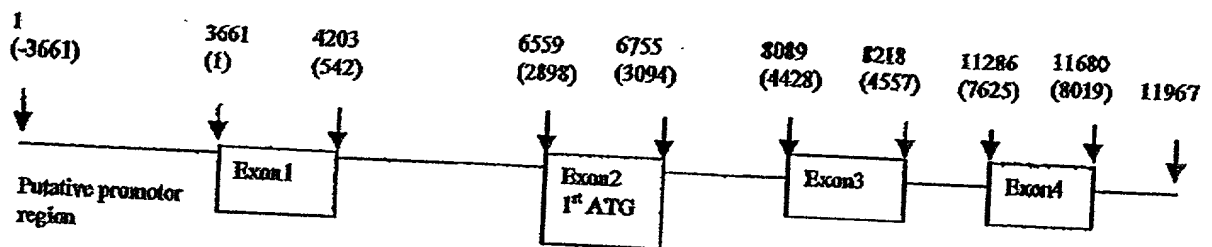


Figure 10

cdna_rat	1	ogagagacgacagaaggttaoggctgcgagaagacgacagaagggtccag
genomic_hu	1	-----
cdna_rat	51	aaaaaggaaagtgtctggaggggagtggggacaaaagcagcgaccaagtga
genomic_hu	1	-----
cdna_rat	101	atgtcacttcagtgaactgaggccaggcaaaaogogcggaaggattttgt
genomic_hu	1	-----
cdna_rat	151	gtagcttgggaocctttcatagacactgatgacacgtttaogcaaaatag
genomic_hu	1	-----
cdna_rat	201	aaatttgaggagaaaogcctgggccttoggaaggagtgattgattagta
genomic_hu	1	-----
cdna_rat	251	cttgcaagtttaggtgactttaaggagaaactaactaatgtatactattga
genomic_hu	1	-----
cdna_rat	301	gggaggaggaagagcattacagagtttccagcagcagcaggaaagcttg
genomic_hu	1	-----
cdna_rat	351	gttaatttggaatggatgatagcattaaaataacagaagcgcctccagg
genomic_hu	1	-----
cdna_rat	401	tctctgaagottoagtcctccagctgaaagocagaaaagactaagccac
genomic_hu	1	-----
cdna_rat	451	taagccttttgatccctttggaagcaaaagaactttccttcctggggtga
genomic_hu	1	-----
cdna_rat	501	agactctcctcagaagatttcctgtctctgcctatgttacaagaggaatc
genomic_hu	1	-----
cdna_rat	551	aaaaccaagacagaagagctcaggatgcagggtgagaggcagggaagtca
genomic_hu	1	-----
cdna_rat	601	oggcttggtgatctccctcaotgctgtctgctggtggtcaocccotggga
genomic_hu	1	-----
cdna_rat	651	gcagggcctgtctctgcgcgtgtgctgtctatgtgccacagagggtgcac
genomic_hu	1	-----
cdna_rat	701	tgtacatttcggtaacctgacctccatccagatggcatcccgccaatgt
genomic_hu	1	-----
cdna_rat	751	ggaacgaataaatttaggatataacagccttactagattgacagaaaacg
genomic_hu	1	-----
cdna_rat	801	actttgatggcctgagcaaacctggagttactcatgtgcacagtaatggc
genomic_hu	1	-----
cdna_rat	851	attcacagagtcagtgacaagaccttctcgggcttgcagtccttgcaagt
genomic_hu	1	-----
cdna_rat	901	ottaaaaatgagctataacaaagtccaaatcattoggaaggatactttct
genomic_hu	1	-----

cdna\_rat genomic\_hu

cDNA_rat	951	acggactcgggagcttggtccggttgacctggatcacaacaacattgaa
genomic_hu	1	-----
cDNA_rat	1001	ttcatcaaccctgaggccttttatggacttacctcgctccgcttggtaca
genomic_hu	1	-----
cDNA_rat	1051	tttagaaggaaaccggctcacaagctccatccagacacatttgtotcat
genomic_hu	1	-----
cDNA_rat	1101	taagotatctccagatatttaaacctctttcattaagtacctgttcttg
genomic_hu	1	-----
cDNA_rat	1151	tctgataacttctgacctccctcccaaaagaaatggtotcotacatgcc
genomic_hu	1	-----
cDNA_rat	1201	aaacotagaaagcctgtatttgcatggaaacccatggacctgtgactgcc
genomic_hu	1	-----
cDNA_rat	1251	atttaaagtgggtgtgtgagtgatgcagggaacccagatataataaaa
genomic_hu	1	-----
cDNA_rat	1301	tgcaagaagacagaagctcttcagtcctcagcaatgtccctttgcat
genomic_hu	1	-----
cDNA_rat	1351	gaacccaggatctctaaaggcagacccctttgctatggtaocatotggag
genomic_hu	1	-----
cDNA_rat	1401	ctttcctatgtacaaagccaaccattgatccatcactgaagtoaaagago
genomic_hu	1	-----
cDNA_rat	1451	ctggttactcaggaggacaatggatctgcctccacctcacctcaagattt
genomic_hu	1	-----
cDNA_rat	1501	catagaacccctttggctccttgtcttttgaacatgacanantntctgyaa
genomic_hu	1	-----
cDNA_rat	1551	ataaggccgacatggtctgtagtatocaaaagccatcaaggacatcacca
genomic_hu	1	-----
cDNA_rat	1601	actgcattcactgaagaaaatgactacatcatgotaaatgogtcattttc
genomic_hu	1	-----
cDNA_rat	1651	cacaaatcttgtgtgcagtgtagattataatcacatccagccagtggtggc
genomic_hu	1	-----
cDNA_rat	1701	aaottctggtttatacagtgactctcctotgatactagaaaggaagccc
genomic_hu	1	-----
cDNA_rat	1751	cagettaccgagactccttcactgtcttctagatataaacaggtggotct
genomic_hu	1	-----
cDNA_rat	1801	taggctgaagacattttaccagcatagaggctgatgtcagagcagacc
genomic_hu	1	-----
cDNA_rat	1851	ctttttggttccaacaagaaaaattgtcttgagotgaaacagaactgcc
genomic_hu	1	-----
cDNA_rat	1901	accacacttagocattacagatocagttttccactgatgtcacaatcgc
genomic_hu	1	-----
cDNA_rat	1951	tttaccaagggcggagatgagagcggagagactcaaattggaocatgatcc
genomic_hu	1	-----

cDNA_rat	2001	tgatgatgaacaatocaaaactggaaogcactgtoctggttggcggcact
genomic_hu	1	-----
cDNA_rat	2051	attgccotgagctgtocaggoaaaggcgacccttcacctcacttggaaatg
genomic_hu	1	-----
cDNA_rat	2101	gottotagctgatgggagtaaaagtgaagagccoccttacgttagcgaggatg
genomic_hu	1	-----
cDNA_rat	2151	ggcgaatcctaataagacaaaaatgggaagttggaactgcagatggctgac
genomic_hu	1	-----
cDNA_rat	2201	agctttgatgcaggtotttaccactgcataagcaccaatgatgcagatgc
genomic_hu	1	-----
cDNA_rat	2251	ggatgttctcacatacaggataactgtggtagagccotatggagaaagca
genomic_hu	1	-----
cDNA_rat	2301	cacatgacagtggagtcagcacacagtggttacgggtgagacgctcgac
genomic_hu	1	-----
cDNA_rat	2351	cttccatgcctttccacgggtgttccagatgottotattagctggattct
genomic_hu	1	-----
cDNA_rat	2401	tccagggaacactgtgttctctcagoccatcaagagacaggcaaattctta
genomic_hu	1	-----
cDNA_rat	2451	acaatgggaccttaagaatattacaggttacgccaaaagatcaaggtcat
genomic_hu	1	-----
cDNA_rat	2501	taccaatgtgtggctgccaccocatcaggggcgacttttccagttttaa
genomic_hu	1	-----
cDNA_rat	2551	agtttcagttcaaaagaaaggccaaaggatggttagcatgacaggagg
genomic_hu	1	-----
cDNA_rat	2601	caggtggatctggacttggagaaocaaactccagtgtttcccttaagcag
genomic_hu	1	-----
cDNA_rat	2651	ccagcatctttgaaactctctgcacagctttgacagggtcagaggctgg
genomic_hu	1	-----
cDNA_rat	2701	aaaacaagtctccggtgtacataggaagaacaaacatagagacttaatac
genomic_hu	1	-----
cDNA_rat	2751	atcggcggcgtggggattccacgctccggcgattcaggagcataggagg
genomic_hu	1	-----
cDNA_rat	2801	cagctccctctctctgctcggagaattgacccgcaacgctgggcagcact
genomic_hu	1	-----
cDNA_rat	2851	tctagaaaaagccaaaaagaattctgtgccaaaaagcaagaaaatacca
genomic_hu	1	-----
cDNA_rat	2901	cagtaaagccagtgccactgggtgttccctcgtggaactcactgaogag
genomic_hu	1	-----
cDNA_rat	2951	gaaaaggatgcctctggcatgattcctcagatgaagaattcatggttct
genomic_hu	1	-----



cDNA_rat	3001	gaaaactaaggcttctggtgtccaggaaggccaccaactgctgactotg
genomic_hu	1	-----
cDNA_rat	3051	gaccagtaaatacatgggttttatgaagagtatagcttctggcacagaagtc
genomic_hu	1	-----
cDNA_rat	3101	tcaactgtgaatccacaacactacaatctgagcaccttootgatttcaa
genomic_hu	1	-----
cDNA_rat	3151	attattttagtgtaacaaacgggtacagctgtgacaaagagtatgaacccat
genomic_hu	1	-----
cDNA_rat	3201	ccatagcaagcaaaatagaagatacaacccaacccaaatcattatc
genomic_hu	1	-----
cDNA_rat	3251	tttccatcagtagctgaaattcgagattctgtctcaggcagggaagagcgc
genomic_hu	1	-----
cDNA_rat	3301	ttcccaaagtgcacacccctgtaacagggggaacatggctacctatggcc
genomic_hu	1	-----
cDNA_rat	3351	ataccaacacatatagtagctttaccagcaaagccagtacagtccttgag
genomic_hu	1	-----
cDNA_rat	3401	ccataaatccaacagaaagttatggacctcagatacctattacaggagt
genomic_hu	1	-----
cDNA_rat	3451	cagcagacctagcagtagtgacatctcttctcacactactgcagacccta
genomic_hu	1	-----
cDNA_rat	3501	gcttctccagtcacccttcagggtcacacaccactgcctogtctttattt
genomic_hu	1	-----
cDNA_rat	3551	cacattcctagaaacaacaatacaggtaacttcccttgtccaggcactt
genomic_hu	1	-----
cDNA_rat	3601	gggaagagagaggacaatttgagcagagggagagttaaaaaccacata
genomic_hu	1	-----
cDNA_rat	3651	gaaccccagttotccgacggcatagacacaggactgtgagggcagcaatc
genomic_hu	1	-----
cDNA_rat	3701	aagggaacctgctaacaaaaatgtgagccaagttccagccacagagtacco
genomic_hu	1	-----
cDNA_rat	3751	tgggatgtgccacacatgtccttcgcagaggggctcacagtggtactg
genomic_hu	1	-----
cDNA_rat	3801	cagcactgtcagttccaagttcatcccaagtgccctcccaaaaactaat
genomic_hu	1	-----
cDNA_rat	3851	aatgttggggtoatagcagaagagctctaccactgtggtcaagaaaccaot
genomic_hu	1	-----
cDNA_rat	3901	gttactattttaaggacaaaacaaatgtagatattgagataataacaacca
genomic_hu	1	-----
cDNA_rat	3951	ctacaaaatattccggaggggaaagtaaccaogtgatttoctacggaagca
genomic_hu	1	-----
cDNA_rat	4001	agcatgaottctgtctccaacatctgtatccctggggaaatctcctgtaga
genomic_hu	1	-----
cDNA_rat	4051	caatagtggtcacctgagcatgcctgggaccatccaaactgggaaagatt
genomic_hu	1	-----
cDNA_rat	4101	cagtggaacaacaccacttccagccccctcagcacaccctcaatacca
genomic_hu	1	-----

cDNA_rat	4151	acaagcacaaaattctcaagaggaaaaactccottgcaccagatctttgt
genomic_hu	1	-----
cDNA_rat	4201	aaataaccagaagaaggaggggatgttaagaatccatatcaattoggtt
genomic_hu	1	-----
cDNA_rat	4251	tacaaagaaccagccgcaagottccaaaatagctcctottttaccc
genomic_hu	1	-----
cDNA_rat	4301	acaggtcagagttccccctcagattctacaactctottgacaagtcggcc
genomic_hu	1	-----
cDNA_rat	4351	accagctctgtctacaacaatggotgccactcagaacaagggcactgaag
genomic_hu	1	-----
cDNA_rat	4401	tagtatcagggtccagaagtctctcagcaggaagaagcagcccttcacc
genomic_hu	1	-----
cDNA_rat	4451	aaactcctotccagtgttccctagcaccataagcaagagatctaatacatt
genomic_hu	1	-----
cDNA_rat	4501	aaacttcttgtcaacggaaacccccacagtgacaagtctactgtactg
genomic_hu	1	-----
cDNA_rat	4551	catctgtcattatgtctgaacccaacgaacaagatccaagaagcaaaa
genomic_hu	1	-----
cDNA_rat	4601	gaccaaataaaggggcctcggaagaacagaaacaacgcaaacaccacccc
genomic_hu	1	-----
cDNA_rat	4651	caggcagggtttctggctatagtgcatactcagctctaacaacagotgata
genomic_hu	1	-----
cDNA_rat	4701	cccccttggctttcagtcattccccacgacaagatgatggtggaaatgta
genomic_hu	1	-----
cDNA_rat	4751	agtgcagttgcttatcaactcaacaacotctcttcttggccataactgaact
genomic_hu	1	-----
cDNA_rat	4801	gtttgagaagtaoaccagacttttgggaataacaacagotttggaacaa
genomic_hu	1	-----
cDNA_rat	4851	cgttgttgagcaaatcacaggagagtaccacagtgaagagcctcagac
genomic_hu	1	-----
cDNA_rat	4901	acaccaccaccactcctcagcagtggtggcgccccagtgcccactccttc
genomic_hu	1	-----
cDNA_rat	4951	cccacctccttttactaaggggtgtggttacagacagcaaagtcacatcag
genomic_hu	1	-----
cDNA_rat	5001	ctttccagatgaogtcaaatagagtgggtcaccatatatgaatcttcaagg
genomic_hu	1	-----
cDNA_rat	5051	cacaatacagatctgcagcaaccctcagcagagggtagccccaatctga
genomic_hu	1	-----
cDNA_rat	5101	gatcataactggaaccactgactctccctctaactctgtttccatccactt
genomic_hu	1	-----
cDNA_rat	5151	ctgtgcagcactaagggtagataaacacagaattctaaatggaagcco
genomic_hu	1	-----
cDNA_rat	5201	tctccctggccagaacacaaatatcagctcaagtcatactccgaaaccat
genomic_hu	1	-----

cDNA_rat genomic_hu	5251 1	tgagaaggcgaagggccagcagtaagcatgtccccccacctcagccttc -----
cDNA_rat genomic_hu	5301 1	cagaggccagcactcatgcctccactggaatacacagaagcatgcagaa -----
cDNA_rat genomic_hu	5351 1	aagagtgtttttgataagaaacctgggtcaaaacccaacttccaaacatct -----
cDNA_rat genomic_hu	5401 1	gccttacgtctctctacctaagactctattgaaaaagccaagaataattg -----
cDNA_rat genomic_hu	5451 1	gaggaaagggtgcaagctttacagttccagctaattcagaogtttttctt -----
cDNA_rat genomic_hu	5501 1	ccttgtgagggtgttgagaccctactgcccatcatccactggaccagagt -----
cDNA_rat genomic_hu	5551 1	ttcatcagganttgaaatatcccaaggacacagaaaagcoggttccacg -----
cDNA_rat genomic_hu	5601 1	tgcttcccaatggcaccttgctccatccagagggtcagtattcaggaccgt -----
cDNA_rat genomic_hu	5651 1	ggacagtaacctgtgtctgtcatttaatacactgggcgtagaccattttca -----
cDNA_rat genomic_hu	5701 1	tgtotctttgtctgtgtgtttttaccggcgaaggattttggacagacatg -----
cDNA_rat genomic_hu	5751 1	tcaaggagatccacagttcactttggaagtactgtggaactaaagtgcaga -----
cDNA_rat genomic_hu	5801 1	gtggagggtatgcccaggcctacggtttccctggatacttgcaaaccaaac -----
cDNA_rat genomic_hu	5851 1	ggtggtctcagaaacggccaagggaagcagaaggtctgggtaaacctg -----
cDNA_rat genomic_hu	5901 1	atggaacattgatcatotataatctgagtcctttatgatcgtgggttttac -----
cDNA_rat genomic_hu	5951 1	aagtgtgtggccagcaaccatctggccaggattcactgttggttaagat -----
cDNA_rat genomic_hu	6001 1	acaagtcatcacagctcccccgtcattatagagcaaaagaggcaagcca -----
cDNA_rat genomic_hu	6051 1	tcgttggggttttagggtggaagtttgaaactgcactgcaaaagga -----
cDNA_rat genomic_hu	6101 1	actccccagcctagtgttccactgggtccctttatgatgggactgaactaaa -----
cDNA_rat genomic_hu	6151 1	accattgcagttgactcattccagattttcttgatccaaatggaactc -----
cDNA_rat genomic_hu	6201 1	tgtatataagaagcatcgctccttcagtgaggggcacttatgagtgcatt -----
cDNA_rat genomic_hu	6251 1	gccaccagctcctcaggtcagagagaagggtagtgattcttactgtgga -----

rat	7351	CttTGGAAGCAGAGTCACGGTCCATCCAAATGGAACCTTGGAgatgagga
le_hu	890	CtaTGGAAGCAGAATCACAGTCCATAAAATGGAACCTTGGAsattagga
rat	7401	acatccGGCTTTCTGACTCTGCGGACTTCACCTGTGTGGttoggagcgag
le_hu	940	atgtgaGGCTTTCAGATTGAGCCGACTTTATCTGTGTGGcccgaaatgaa
rat	7451	ggaGGAGAGAGTGTGTGGTAGTGCACTTAGAAGTCCTAGAAATGCTGAG
le_hu	990	ggtGGAGAGAGCGTGTGGTAGTACAGTTAGAAGTACTGGAATGCTGAG
rat	7501	AAGACCAACATTTCAGAAACCCATTCAACGAAAagtcacgcgccaagctg
le_hu	1040	AAGACCGACATTTAGAAATCCATTTAATGAAAAaatagttgccagctgg
rat	7551	gcaagcccgtagCACTGAAGTGTCTGTGGATGGGAACCCCCACCTGAA
le_hu	1090	gaaagtcacacAGCATTGAATTGCTCTGTGATGGTAACCCACCACCTGAA
rat	7601	attacCTGGATCTTACCTGACGGCACACAGTTTGCTAACAGACCacacaa
le_hu	1140	ataatCTGGATTTTACCAAATGGGCACACGATTTTCCAATGGACCAcaaaag
rat	7651	ttccccGTATCTGATGGCAGGCAATGGCTCTCTCATCTTTACAAAgcaa
le_hu	1190	ttatcaGTATCTGATAGCAAGCAATGGTTCTTTTATCATTTCTAAAacaa
rat	7701	ctcggaacaaagtCAGGGAAGTATCGCTGTGCAGCCAGGAATAAGGTTGGC
le_hu	1240	ctcgggaggatgCAGGAAATATCGCTGTGCAGCTAGGAATAAAGTTGGC
rat	7751	TAcacgcagaaactcatcctgttagagATTGGGCAGAAGCCAGTCATTCT
le_hu	1290	TAtattgagaaattagtoatattagaaATTGGCCAGAAGCCAGTTATTCT
rat	7801	GACATACGAACCCAGGGatggtgaagagcgtCAGTGGGGAACCGTTATCAC
le_hu	1340	TACCTATGCACCAGGCAcagtaaaaggoatCAGTGGAGAATCTCTATCAC
rat	7851	TGCATTGTGTGTCTGATGGgATCCCCAAGCCAAATGTCAAGTGGACTACA
le_hu	1390	TGCATTGTGTGTCTGATGGaATCCCTAAGCCAAATATCAAATGGACTATG
rat	7901	CCGGGTGGccATGTAATCGACAGGCCTCAAGTGGATGGAAAATACATAcT
le_hu	1440	CCAAAGTGGtTATGTAGTAGACAGGCCTCAAATTAATGGGAAATACATaT
rat	7951	GCATGAAAATGGCAGCTGGTCATCAAAGCAACAACAGCtccagaccaaAG
le_hu	1490	GCATGACAATGGCACCTTAGTCATTAAAGAAGCAACAGCttatgacagAG
rat	8001	GAAATTATATCTGTAGGGCTCAAACAGTGTGGCCAggcagttattago
le_hu	1540	GAAACTATATCTGTAAGGCTCAAATAGTGTGGTCatacactgattact
rat	8051	gtgtCAGTGATGGTTGTGGCCTACCCTCCCCGAATCATAAActacctACC
le_hu	1590	gttcCAGTAATGATTGTAGCCTACCCTCCCCGAATTACAAatcgtccACC
rat	8101	CAGGAACATGCTCAGGAGGACAGGGGAAGCCatgCAGCTCCACTGTGTGG
le_hu	1640	CAGGAGTATTGTCACCAGGACAGGGGCAGCctttCAGCTCCACTGTGTGG
rat	8151	CCTTGGGAATCCCCAAGCCaaAAGTCACCTGGGAGACGCCAAGACACTCC
le_hu	1690	CCTTGGGAGTTCCCAAGCCagAAATCACATGGGAGATGCCCTGACCACTCC
rat	8201	CTGCTCTCAAaagcaacagcaagaaaaacccCATAGAAGTGAGATGCTTCA
le_hu	1740	CTTCTCTCAAcggcaagtaagagaggacaCATGGAACTGAGCAGCTTCA
rat	8251	CCCACAAGGTACGCTgGTCATTTCAGAATCTCCAAACCTCGGATTCCGGag
le_hu	1790	CTTACAAGGTACCCTaGTCATTTCAGAATCCCAAACCTCCGATTCTGGga
rat	8301	tcTATAAGTGACAGAGCTCAGAACCTACTTGGgacTGATTACGCAACAACT
le_hu	1840	taTACAAATGCACAGCAAAGAACCCTTGGtagTGATTATGCAGCAACG
rat	8351	TACATCCAGGTACTCTGACAGGAagggggagactaaaaattoaacagaagt
le_hu	1890	TATATTCAAGTAATCTGACATGAAataataaagtcaacaacatctgggca

**Figure 12**

ATGAAGGTAAAGGCAGAGGAATCACCTGCTTGCTGGTCTCCTTTGCTGTGATCTGCCTGGTCGCCACC  
CCTGGGGGCAAGGCCTGTCTCGCCGCTGTGCCTGTTATATGCCCTACGGAGGTACACTGCACATTTTCGG  
TACCTGACTTCCATCCCAGACAGCATCCCGCCCAATGTGGAACGCATCAATTTAGGATACAACAGCTTG  
GTTAGATTGATGGAAACAGATTTTCTGGCCTGACCAAACCTGGAGTTACTCATGCTTCACAGCAATGGC  
ATTCACACAATCCCTGCACAAGACCTTCTCAGATTGGCAGGCCCTTGACAGTCTTAAAAATGAGCTATAAT  
AAAGTCCGAAAACTTCAGAAAAGATACTTTTTATGGCCTCAGGAGCTTGACACGATTGCACATGGACCAC  
AACAAATATTGAGTTTATAAAACCAGAGGTTTTTATGGGCTCAACTTTCTCCGCCTGGTGCACCTTGGAA  
GGAAATCAGCTCACTAAGCTCCACCCAGATACATTTGTCTCTTTGAGCTACCTCCAGATATTTAAAAATC  
TCTTTTCATTAAGTTCTTATACTTGTCTGATAACTTCTTGACCTCCCTCCCTCAAGAGATGGTCTCCTAT  
ATGCCTGACCTAGACAGCCTTTACCTGCATGGAAACCCATGGACCTGTGATTGCCATTTAAAGTGGTTG  
TCTGACTGGATACAGCCAGATGTAATAAAATGCAAAAAGATAGAAGTCCCTCTAGTGCTCAGCAGTGT  
CCACTTTGCATGAACCCCTAGGACTTCTAAAGGCAAGCCGTTAGCTATGGTCTCAGCTGCAGCTTTCCAG  
TGTGCCAAGCCAACCATTGACTCATCCCTGAAATCAAAGAGCCTGACTATTCTGGAAGACAGTAGTTCT  
GCTTTCATCTCTCCCAAGGTTTCATGGCACCCTTTGGCTCCCTCACTTTGAATATGACAGATCAGTCT  
GGAAATGAAGCTAACATGGTCTGCAGTATTCAAAGGCCCTCAAGGACATCACCCATTGCATTCACTGAA  
GAAAATGACTACATCGTGCTAAATACTTCATTTTCAACATTTTGGTGTGCAACATAGATTACGGTCAC  
ATTGAGCCAGTGTGGCAAATTTTGGCTTTGTACAGTGATTCTCCTCTGATACTAGAAAGGAGCCACTTG  
CTTAGTGAAACACCGCAGCTCTATTACAAATATAAACAGGTGGCTCCTAAGCCTGAAGACATTTTACC  
AACATAGAGCAGATCTCAGAGCAGATCCCTCTTGGTTAATGCAAGACCAAATTTCTTGCACTGAAC  
AGAATGCCACCACATTCAGTACATTACAGATCCAGTACTCCAGTGATGCTCAAATCACTTTACCAAGA  
GCAGAGATGAGGCCAGTGAAACACAAATGGACTATGATTTCAAGGGATAACAATACTAAGCTGGAACAT  
ACTGTCTTGGTAGGTGGAACCGTTGGCCTGAACTGCCCAGGCCAAGGAGACCCCCACACGTTGGAT  
TGGCTTCTAGCTGATGGAAGTAAAGTGAGAGCCCTTATGTCACTGAGGATGGACGGATCCTAATAGAC  
AAAAGTGGAAAATTGGAATCCAGATGGCTGATAGTTTGGACACAGGCGTATATCACTGTATAAGCAGC  
AATTATGATGATGCAGATATTCTCACCTATAGGATAACTGTGGTAGAACCTTTGGTCAAGCCTATCAG  
GAAAATGGGATTTCATCACACAGTTTTCATTGGTGAAACACTTGATCTTCCATGCCATTCTACTGGTATC  
CCAGATGCCTCTATTAGCTGGGTTATTCCAGGAAACAATGTGCTCTATCAGTCATCAAGAGACAAGAAA  
GTTCTAAACAATGGCACATTAAGAATATTACAGGTCAACCCGAAAGACCAAGGTTATTATCGCTGTGTG  
GCAGCCAACCCATCAGGGGTTGATTTTTTGTATTTTCCAAGTTTCAGTCAAGATGAAAGGACAAAGGCC  
TTGGAGCATGATGGAGAAACAGAGGGATCTGGACTTGATGAGTCCAATCCTATTGCTCATCTTAAGGAG  
CCACCAGGTGCACAACCTCCGTACATCTGCTCTGATGGAGGCTGAGGTTGGAAAACACACCTCAAGCACA  
AGTAAGAGGCACAACATATCGGAATTAACACTCCAGCGACGTGGAGATTCAACACATCGACGTTTTAGG  
GAGAATTAGGAGGCATTTCCCTCCCTCTGCTAGGAGAATTGACCCACAACATTTGGGCGGCAGTGTGGAG  
AAAGCTAAAAAGAATGTATGCCAGACAAGCGAGAAAAATACCACAGTGAGCCACCCCAAGTGTGCTCACC  
CAACTCCCAACATACCTGGTGAAGAAGACGATTCTCAGGCATGCTCTGCTCTACATGAGGAATTTATG  
GTCCCGGCCACTAAAGCTTTGAACCTTCCAGCAAGGACAGTGACTGCTGACTCCAGAACAATATCTGAT  
AGTCTATGACAAACATAAAATATGGCACAGAATTCTCTCCTGTTGTGAATTCACAAATACTACCACCT  
GAAGAACCCACAGATTTCAAACCTGTCTACTGCTATTAAACTACAGCCATGTCAAAGAATATAAACCCA  
ACCATGTCAAGCCAAATACAAGGCACAACCAATCAACATTCACTGCTTTTCCACTGCTACTTGGGA  
GCAACTGAATTTTCAAGACTCTGACCAGATGGGAAGAGGAAGAGCATTTCCAAAGTAGACCCCAATA  
ACAGTAAGGACTATGATCAAAGATGTCAATGTCAAATGTCTAGTAGCACCACCAACAACTATTATTA  
GAGTCAGTAAATACCACAAATAGTCATCAGACATCTGTAAGAGAAGTGAGTGAACCCAGGCACAATCAC  
TTCTATTCTCACACTACTCAAATACTTAGCACCTCCACGTTCCCTTCAGATCCACACACAGCTGCTCAT  
TCTCAGTTTCCGATCCCTAGAAATAGTACAGTTAAATCCCGCTGTTTCAGACGCTTTGGGAGGCAGAGG  
AAAATTGGCGGAAGGGGGCGGATTATCAGCCCATATAGAACTCCAGTTCTGCGACGGCATAGATACAGC  
ATTTTCAGTCAACAACCAAGAGGTTCTTCTGAAAAAGCACTACTGCATTCTCAGCCACAGTGCTCAAT  
GTGACATGTCTGTCTCTTCTCCAGGGAGAGGCTCACCCTGCCACAGCAGCATTGTCTTTTCCAAGT  
GCTGCTCCCATCACCTTCCCAAGAGGATGACATTGTAGAGTCCCATCAGAAGAGTCTACAACCTTAGTC  
CAGAATCCACTATTACTACTTGAGAACAACCCAGTGTAGAGAAAAACAACCCACAATAAAATATTTTC  
AGGACTGAAATTTCCCAAGTGACTCCAACCTGGTGCAGTCATGACATATGCTCCAACATCCATACCCATG  
GAAAAAATCACAAGTAAACGCCAGTTACCCACGTGTGTCTAGCACCATGAAGCTAAAAGAGATTCA  
GTGATTACATCGTCACTTTTCAAGTGCTATCACCAGCCACCAATGACTATTATAGCCATTACAAGGTTT  
TCAAGAAGGAAAATTTCCCTGGCAACAGAACTTTGTAAATAACCATAACCCAAAAGGCAGATTAAGGAAT  
CAACATAAAGTTAGTTTACAAAAAGCACAGCTGTGATGCTTCTTAAACATCTCCTGCTTTACCACAG  
AGACAAAGTTCCCTTTTCATTTTCAACCACTTTCAACAAGTGTGATGCAATTCATCTAATACCTTG  
ACTACCGCTCACCACACTACGACCAAAACACACAATCCTGGAAGTCTTCCAACAAAGAGGAGCTTCCC  
TTCCACCCCTTAACCTTATGCTTCTTAGTATTATAAGCAAAGACTCAAGTACAAAAAGCATCATATCA  
ACGCAAAACAGCAATACCAGCAACAACCTCTACCTTCCCTGCATCTGTCTCATCTTATGAAACCCAAACA  
GAGAGATCTAGAGCACAAACAATACAAAGAGAACAGGAGCCTCAAAGAAGAAGAGGACTGACCCAAAC  
ATCTCTCCAGACCAGAGTTCTGGCTTCACTACACCCACTGCTATGACACCTCTGCTCTGGCATTCACT

CATTCCCCACCAGAAAACACAAC TGGGATTTCAAGCACAATCAGTTTTTCATTCAAGAACTCTTAATCTG  
ACAGATGTGATTGAAGAACTAGCCCAAGCAAGTACTCAGACTTTGAAGAGCACAATTGCTTCTGAAACA  
ACTTTGTTCAGCAAATCACACCAGAGTACCACAAC TAGGAAAGCATCATTAGACACTCCCATACCACCA  
TTCTTGAGCAGCAGTGTACTCTAATGCCAGTTC CATCTCCCCCTCCCTTTACTCAGAGAGCAGTTACT  
GACACACGTGGCGACTCCCATTTCCGGCTTATGACAAATACAGTGGTCAAGCTGCACGAATCCTCAAGG  
CACAATCTCCAAATGCCAAGTTCACAATTGGAACCACTCACTTCATCTACCTCTAATCTGTTACATTCT  
ACTCCCATGCCAGCACTAACAACAGTTAAATCACAGAATTCCAAATTAAC TCCATCTCCCTGGGCAGAA  
TACCAATTTTGGCACAACCATACTCAGACATTGCTGAAAAAGGCCAAAAAGCCAGAAGTAAGCATGTTG  
GCTACTACAGGCCTGTCCGAGGCCACCCTCTTGT TTTAGATTGGGATGGACAGAAGAACACAAAGAAG  
AGTGACTTTGATAAGAAACCAGTTC AAGAAGCAACAAC TCCAAACTCCTTCCCTTTGACTCTTTGTCT  
AGGTATATATTTGAAAAGCCCAGGATAGTTGGAGGAAAAGCTGCAAGTTT TACTATTCCAGCTAACTCA  
GATGCCTTTCTTCCCTGTGAAGCTGTTGGAAATCCCCTGCCACCATTCA TTGGACCAGAGTTTCAGGA  
CTTGATTTATCTAGAGGAAACCAGAATAGCAGGGTCCAGGTTCTCCCCAATGGTACCCTGTCCATCCAG  
AGGGTGGAATTCAGGACCGCGACAGTACTTGTGTTCCGCATCCAATCTG TTTGGCACAGACCACCTT  
CATGTACCTTGTCTGTGGTTTCTTATCTCTCC CAGGATCCTGGAGAGACGTACCAAAGAGATCACAGTT  
CATTCCGGAAGCACTGTGGAAC TGAAGTGCAGAGCAGAAGGTAGGCCAAGCCCTACAGTTACCTGGATT  
CTTGCAAAACCAACAGTTGTCTCAGAATCATCC CAGGGAAGTAGGCAGGCTGTGGTGACGGTTGACGGA  
ACATTGGTCTCCACAATCTCAGTATTTATGACC GTGGCTTTTACAAATGTGTGGCCAGCAACCCAGGT  
GGCCAGGATTCAGTGTGGTTAAATACAAAGTCA TTGCAGCACCACCTGTTATTCTAGAGCAAAGGAGG  
CAAGTCATTGTAGGCACTGGGGTGAAAGTTTAA AACTGCCCTGTACTGCAAAAGGAACTCCTCAGCCC  
AGCGTTTACTGGGTCTCTCTGATGGCACTGAAG TGAAACCATTACAGTTTACCAATTCCAAGTGTTC  
TTATTTTCAAATGGGACTTTGTATATAAGAAAC CTAGCCTCTTCAGACAGGGGCACTTATGAATGCATT  
GCTACCAGTTCCACTGGTTCGGAGCGAAGAGTA GTAATGCTTACAATGGAAGAGCGAGTGACCAGCCCC  
AGGATAGAAGCTGCATCCCAGAAAAGGACTGAAG TGAATTTTGGGGACAAATTA TACTACTGAAC TGCTCA  
GCCACTGGGGAGCCCAACCCCAAATAATGTGGAG GTTACCATCCAAGGCTGTGGTTCGACCAGTGGAGC  
TGGATCCACGTCTACCCTAATGGATCCCTGTTT ATTTGGATCAGTAACAGAAAAAGACAGTGGTGTCTAC  
TTGTGTGTGGCAAGAAACAAATGGGGGATGATCT GATACTGATGCATGTTAGCCTAAGACTGAAACCT  
GCCAAAATTGACCACAAGCAGTATTTTAGAAAAG CAAGTGCTCCATGGGAAAGATTTCCAAGTAGATTGC  
AAAGCTTCCGGCTCCCCAGTGCCAGAGATATCT TGGAGTTTGCCTGATGGAACCATGATCAACAATGCA  
ATGCAAGCCGATGACAGTGGCCACAGGACTAGGA GATATACCCTTTTCAACAATGGAAC TTTATACTTC  
AACAAAGTTGGGGTAGCGGAGGAAGGAGATTATA CTTGCTATGCCCAGAACACCTTAGGGAAAGATGAA  
ATGAAGGTCCACTTAACAGTTATAACAGCTGCTC CCCGGATAAGGCAGAGTAACAAAACCAACAAGAGA  
ATCAAAGCTGGAGACACAGCTGTCTTGACTGTG AGGTCACTGGGGATCCCAAACCAAAAATATTTTGG  
TTGCTGCCTTCCAATGACATGATTTCTTCTCTC CATTGATAGGTACACATTT CATGCCAATGGGTCTTTG  
ACCATCAACAAAGTGAAACTGCTCGATTCCTGG AGAGTACGTATGTGTAGCCCGAAATCCCAGTGGGGAT  
GACACCAAAATGTACAAACTGGATGTGGTCTCTA AAACCTCCATTAAATCAATGGTCTGTATACAAACAGA  
ACTGTTATTAAAGCCACAGCTGTGAGACATTCCA AAAAAACACTTTGACTGCAGAGCTGAAGGGACACCA  
TCTCCTGAAGTCATGTGGATCATGCCAGACAATA TTTTCTCACAGCCCCATACTATGGAAGCAGAATC  
ACAGTCCATAAAAATGGAACCTTGGAATTAGGAAT GTGAGGCTTTCAGATT CAGCCGACTTTATCTGT  
GTGGCCCGAAATGAAGGTGGAGAGACGCTGTTG GTTAGTACAGTTAGAAGTACTGGAAATGCTGAGAAGA  
CCGACATTTAGAAATCCATTTAATGAAAAAATAG TTGCCAGCTGGGAAAGTCCACAGCATTGAATTGC  
TCTGTTGATGGTAACCCACCACCTGAAATAATCT GGATTTTACC AAATGGCACACGATTTTCCAATGGA  
CCACAAAGTTATCAGTATCTGATAGCAAGCAATG GTTCTTTTATCATTTCTAAAACAAC TCGGGAGGAT  
GCAGGAAAATATCGCTGTGCAGCTAGGAATAAAG TTGGCTATATTGAGAAATTAGTCATATTAGAAATT  
GGCCAGAAGCCAGTTATTCTTACCTATGCACCAG GGACAGTAAAAAGGCATCAGTGGAGAATCTCTATCA  
CTGACATTGTGTGTCTGATGGAATCCCTAAGCC AAATATCAAATGGACTATGCCAAGTGGTTATGTAGTA  
GACAGGCCCTCAAATTAATGGGAAATACATATTGC ATGACAATGGCACCTTAGTCATTAAAGAAGCAACA  
GCTTATGACAGAGGAAACTATATCTGTAAGGCTC AAAATAGTGTGGTCATACACTGATTACTGTTCCA  
GTAATGATTGTAGCCTACCCTCCCCGAATTACA AATCGTCCACCCAGGAGTATTGTCAACAGGACAGGG  
GCAGCCTTTCAGCTCCACTGTGTGGCCTTGGGAG TTCCCAAGCCAGAAATCACATGGGAGATGCCTGAC  
CACTCCCTTCTCTCAACGGCAAGTAAAGAGAGGAC ACATGGAAGTGAGCAGCTTCACTTACAAGGTACC  
CTAGTCATT CAGAATCCCCAAACCTCCGATTCT GGGATATACAAATGCACAGCAAAGAACCCACTTGGT  
AGTGATTATGCAGCAACGTATATTCAAGTAATCTGA

Figure 13

A

Region			Region Length	% identity	% positives	% gaps
General	Rat	Human				
1-655	1-655	1-653	655	76	86	0
656-726	656-726	654-724	71	46	62	0
727-779	727-779	725-777	53	77	86	0
780-1634	780-1617	778-1612	655	38	53	4
1635-end	1618-end	1613-end	980	74	85	0
	Total	Alignment	2614	62	74	1

B

Region			% identity	% positives	% gaps
Rat	Mouse	Length			
1-238	1-238	238	91	92	1

C

Region			Region Length	% identity
General	Rat	Human		
1-1965	1-1965	1-1965	1965	83
1966-2178	1966-2178	1966-2178	213	86
2179-2337	2179-2337	2179-2337	159	86
2338-4893	2338-4893	2338-4863	2565	63
4894-7833	4852-7791	4864-7761	2940	80
	Total	Alignment	7842	80

D

Region			Region Length	% identity
General	Rat	Mouse		
1-720	1-718	1-720	720	93

Figure 14

rat_cDNA	CGAGAGACGACAGAGGTTACGGCTGCCGAGAACGACAGAGGGTCCAGAAAAGGAAA
human 5+3 corrected	-----CAGAGGGTCCAGGAAA-GGAAA
mus_cDNA_5	
rat_cDNA	GTGCTGGAGGGGAGTGGGACAAAGCAGCGACCAAGTGAATGTCACTTCAGTGAAGT
human 5+3 corrected	-----GTACTGGAGGGGAGTGGGACAAAGCAGCGACCAAGGAACATCGCTTCAGTGAAGT
mus_cDNA_5	
rat_cDNA	GCCAGGCAAAAGCGCGGGAAGGATTTTGTGTAGCTTGGGAOCCTTTTATAGACACTGAT
human 5+3 corrected	-----GCCAGGCAAAAGGAGCGGGAAGGATTTATATGTAGCTTGGGAOCCTTTTATAAACACTGAT
mus_cDNA_5	
rat_cDNA	GACACGTTTACGCAAAATA-GAAATTTGAGGAGAAACGCTGGGCTTCGGAAAG---GA
human 5+3 corrected	-----GACGTTTGTGCAAGGCAAGCAATTTGAGGAGAAACGCTGGGACGTCGGAAAGAAGGA
mus_cDNA_5	
rat_cDNA	GTGATTGATTAGTACTTGCAGTTTAGGTGACTTTAAGGAGACTAACTAATGTATACTA
human 5+3 corrected	-----GTGATCGATTAGTACTTGTAACTTTAGGTGAGTTT---GAGAACTAACTAAGCTATACTA
mus_cDNA_5	
rat_cDNA	TTGAGGGAGGAGGAAAGCATTACAGAGTTTCCAGCAGCAGCAGGAAAGCTTTGGTTAAT
human 5+3 corrected	-----TTGAGGGAGGAGGAAAGCATT-----CCAGCAGCAGCAGGAAAGCTTTGGTTAAT
mus_cDNA_5	
rat_cDNA	TTGGAATGGATGATAGCAATTAATAACAGAAAGCGOCTCCAGGTCTGTGAGCTTCAGT
human 5+3 corrected	-----TTGGAATGTATGATACCAATTAATAACAGAAAGCGOCTCCAGTTCTGTGAGAGTTCAGT
mus_cDNA_5	
rat_cDNA	CCCCAGCTGAAAGCCAGAAAGACTAAGCCCCAAGCCCTTTTGTATCCCTTTGGAGCA
human 5+3 corrected	-----CCCCAGCT-----A-GTGTAAAGCTACTAAGCCCTTTTGTATCCCTTTGGAGCA
mus_cDNA_5	
rat_cDNA	AAGAACTTTCCTTCCCTGGGGTGAAGACTCTCCTCAGAAGATTTCTGTCTCTGCCTATG
human 5+3 corrected	-----AAGAACTTTCCTTCAATCAGGTGAAGGCTCTCCTCAGAAGATTTCTGTCTCTGCCTATG
mus_cDNA_5	
rat_cDNA	TTACAGAGGAATCAAAACCAAGACAGAGAGCTCAGGATCCAGGTGAGAGGCAGGGAAG
human 5+3 corrected	-----TTACAGAGGAATCAAAAGCAAGACAGAGAGCTCAGGATCCAGAGAGAGGCAGGGAAG
mus_cDNA_5	*** ** * ***** *
rat_cDNA	TCAGCGGCTTGTGTCTCCCTCAGTGTCTGTGCTGGTGGTCACCCTGGGAGCAGGG
human 5+3 corrected	-----TCACCTGCTTGTGTCTCCCTTGTGTGTGTGTGCTGGTGGTCACCCTGGGAGCAGGG
mus_cDNA_5	TCAGCTGCTTGTGTCTCCCTCAGTGTCTGTGCTGGTGGTCACCCTGGGAGCAGGG *** * ***** ** ***** * *** ***** * ***** * *
rat_cDNA	CCTGTCTCTGCGCGCTGTGCTGTATGTGCCCCAGAGGTGCACTGTACATTTGGGTACC
human 5+3 corrected	-----CCTGTCTCTGCGCGCTGTGCTGTATGTGCCCCAGAGGTGCACTGTACATTTGGGTACC
mus_cDNA_5	TCTGTCTCTGCGCGGTGTGCTGTATGTGCCCCAGAGGTGCACTGTACATTTGGGTACC ***** ***** *** ***** * ***** ***** *
rat_cDNA	TGACCTCCATCCAGATGG-CATCCCGGCCAATGTGGAACGAATAAATTTAGGATATAC
human 5+3 corrected	-----TGACTTCCATCCAGACAG-CATCCCGGCCAATGTGGAACGCATCAATTTAGGATATAC
mus_cDNA_5	TGACCTCCATCCAGAGGGGATCCAGCCCAATGTGGAACGAGTCAATTTAGGATATAC **** ***** * ***** ***** * ***** * *



rat_cDNA human 5+3 corrected mus_cDNA_5	AGCCTTACTAGATTGACAGAAAACGACTTTGATGGCCTGAGCAAACTGGASTTACTCATG AGCTTGCTTAGATTGATGAAACAGATTTTCTGGCCTGACCAAACTGGASTTACTCATG AGCCTCACTAGATTGACAGAAAATGACTTTTCTGGCCTGAGCAGACTGGAGTTACTCATG *** * ***** ** * ***** ** *****
rat_cDNA human 5+3 corrected mus_cDNA_5	CTGCACAGTAATGGCATTACAGAGTCAGTGACAAGACCTTCTCGGGCTTGCACTCCTTG CTTCACAGCAATGGCATTACAGCAATTCCTGACAAGACCTTCTCAGATTGCAAGGCTTG CTGCACAGCAATGGCATTACAGAGTCAGTGACAAGACCTTCTCGGGCTTGCACTCCTTG ** ***** ** *****
rat_cDNA human 5+3 corrected mus_cDNA_5	CAGGTCTTAAAAATGAGCTATAACAAAGTCCAAATCATTGGAAGGATACTTTCTACGGA CAGGTCTTAAAAATGAGCTATAACAAAGTCCGAAAACCTCAGAAAGATACTTTTATGGC CAGGTCTTAAAAATGAGCTATAACAAAGTCCAAATCATTGGAAGGATACTTTGTATGGA *****
rat_cDNA human 5+3 corrected mus_cDNA_5	CTCGGAGCTTGGTCCGCTTGACCTGGATCACAACAACATTGAATTCATCAACCTCGAG CTCAGGAGCTTGACAGGATTGCACATGGACCAACAAATATTGAGTTTATAAACCAGAG CTCAGGAGCTTGACCCGCTTGACCTGGATCACAACAACATTGAGTTTATCAACCCGAG *** ***** ** *****
rat_cDNA human 5+3 corrected mus_cDNA_5	GCGTTTATGGACTTACCTCGCTCGCTTGGTACATTAGAGGAAACCGGCTCACAAG GTTTTTATGGGCTCACTTTCTCGCTGGTGCATTTGGAAGGAAATCAGCTCACTAAG GCGTTTATCGGACTCACCTTCTCGCTGGTACATCTAGAGGAAACCGGCTCACAAG * ***** ** * *****
rat_cDNA human 5+3 corrected mus_cDNA_5	CTCCATCCAGACACATTGTCTCATTAGCTATCTCCAGATATTTAAACCTCTTTTCAAT CTCCACCCAGATACATTGTCTCTTTGAGCTACCTCCAGATATTTAAATCTCTTTCAAT CTCCATCCAGACACATTGTCTCTTTGAGCTATCTCCAGATATTTAAACCTCTTTTCAAT *****
rat_cDNA human 5+3 corrected mus_cDNA_5	AAGTACCTGTCTTGTCTGATAACTTCCT-GACCTCCCTCCCAAAAGAAATGGTCTCCTA AAGTCCCTATACTTGTCTGATAACTTCCT-GACCTCCCTCCCAAGAGATGGTCTCCTA AAGNACCTGTACTTGTATGATAACTTCATTGACCTCCCTCCCAAAAGAAATGGTCTCCTC *** * *****
rat_cDNA human 5+3 corrected mus_cDNA_5	CATGCAAACTAGAAAGCCTGTATTGCAATGGAACCCATGGACCTGTGACTGCCATT TATGCTGACCTAGACAGCCTTTAAGCTGATGGAACCCATGGACCTGTGATTGCCATT TATGCCAAACTAGAAAGCCTTTACTTGCATGGAACCCATGGACCTGTGACTGCCATT *****
rat_cDNA human 5+3 corrected mus_cDNA_5	AAAGTGGTTGTCTGAGTGGATGCAAGGAAACCCAGATATAATAAATGCAAGAAAGACAG AAAGTGGTTGTCTGACTGGATACAGGNNNNOCAGATGTAATAAATGCAAAAAAGATAG AAAGTGGTTGTCCGAGTGGATGCAAGGAAACCCAGTA-ACTATCTTGT-TTGTITG *****
rat_cDNA human 5+3 corrected mus_cDNA_5	AAGCTCTTCCAGTCTCAGCAATGTCCCT-TTGCATGAACCCAGGATCTCTAAGGCA AAGTCCCTCTAGTCTCAGCAGTGTCCCT-TTGCATGAACCCAGGACTCTCTAAGGCA TTTCTTTTATARKAGTATTTCCTCAATTTCAATTAAGATGATATCCCAAGATC- * * * * *
rat_cDNA human 5+3 corrected mus_cDNA_5	GACCTTTGCTATGCTACCATCTGAGCTTTCTATGTACAAGCCAAACCATTTGATCCAT AGCCCTTAGCTATGGTCTCAGCTGCAGCTTTCCAGTGTGCCAGCCAAACCATTTGACTCAT -CCCATAACTCCCCCCA-----CTTCCCTACCTACCCATTC-CCATTTTGGC ** * * * * * * * * *
rat_cDNA human 5+3 corrected mus_cDNA_5	CACTGAAGTCAAAAGAGCCTGTTACTCAGGAGGACATGGATCTGCTCCACCTCAGCTC CCCTGAATCAAAAGAGCCTGACTATTCTGGAAGACAGTAGTTCTGCTTCTCTCTCC CCTGGCATTCCCT- * * * *
rat_cDNA human 5+3 corrected mus_cDNA_5	AAGATTTATAGAACCTTTGGCTCTGTCTTTGAACATGACNNANTNTCTGGAATA AAGTTTCATGGAACCTTTGGCTCCCTCACTTTGAATATGACAGATCAGTCTGGAATG -----
rat_cDNA human 5+3 corrected mus_cDNA_5	AGGCGACATGGTCTGTAGTATCCAAAGCCATCAAGGACATCACCAACTGCATTCACTG AAGCTAACATGGTCTGCAATATCCAAAGCCCTCAAGGACATCACCCATTGCATTCACTG -----

rat\_cDNA  
human 5+3 corrected  
mus\_cDNA\_5

AAGAAATGACTACATCATGCTAAATGCGTCATTTTCCACAAATCTTGTGTGAGTGTAG  
AAGAAATGACTACATCGTGCTAAATACTTCATTTTCAACATTTTGGTGTGCAACATAG

rat\_cDNA  
human 5+3 corrected  
mus\_cDNA\_5

ATTATAATCACATCCAGCCAGTGTGGCACTTCTGGCTTTATACAGTGACTCTCTCTGA  
ATTACGGTCACATTCAGCCAGTGTGGCAATTTTGGCTTTGTACAGTGATTCTCTCTGA

rat\_cDNA  
human 5+3 corrected  
mus\_cDNA\_5

TACTAGAAAGGAAGCCCGAGCTTACCGAGACTCCTTCAGTGTCTTCTAGATATAAACAGG  
TACTAGAAAGGAAGCCACTTGCTTAGTGAACACCCGAGCTCTATTACAAATATAACAGG

rat\_cDNA  
human 5+3 corrected  
mus\_cDNA\_5

TGGCTCTTAGGCGTGAAGACATTTTACCAGCATAGAGGCTGATGTGAGAGCAGACCGTT  
TGGCTCCTAAGCGTGAAGACATTTTACCACATAGAGGCAGATCTCAGAGCAGATCCCT

rat\_cDNA  
human 5+3 corrected  
mus\_cDNA\_5

TTTGGTTCCACAAAGAAAAATTGTCTTGCACTGAACAGACTGCCACCACTTAGCA  
CTTGGTTAATGCAGACCAAAATTCCTTGCACTGAACAGACTGCCACCACTTAGCA

rat\_cDNA  
human 5+3 corrected  
mus\_cDNA\_5

CATTACAGATCCAGTTTTCAGTGTGCTCAAATCGCTTTACCAAGGGCGGAGATGAGAG  
CATTACAGATCCAGTACTCCAGTGTGCTCAAATCAGTTTACCAAGAGCAGAGATGAGGC

rat\_cDNA  
human 5+3 corrected  
mus\_cDNA\_5

CGGAGAGACTCAAATGGACCATGATCTGTATGAACAAATCCAAACTGGAAACGCACTG  
CAGTGAACACAAATGGACTATGATTTCAAGGGATAACATACTAAGCTGGAAACATACTG

rat\_cDNA  
human 5+3 corrected  
mus\_cDNA\_5

TCCTGGTTGGCGGCACTATTGCGGTGAGCTGTCCAGGCAAGGCGACCGTTCACTCACT  
TCTTGGTAGGTGGAAACCGTTGGCCTGAAGTCCCAGGCAAGGAGACCCACCCACAGG

rat\_cDNA  
human 5+3 corrected  
mus\_cDNA\_5

TGGAATGGCTTCTAGCTGATGGAGTAAAGTGAGAGCCCGTTACGTTAGGAGGATGGGC  
TGGATTGGCTTCTAGCTGATGGAGTAAAGTGAGAGCCCGTTATGTCACTGAGGATGGAC

rat\_cDNA  
human 5+3 corrected  
mus\_cDNA\_5

GAATCCTAATAGACAAAAATGGGAAGTTGGAAGTGCAGATGGCTGACAGCTTTGATGCAG  
GGATCCTAATAGACAAAAATGGGAAGTTGGAAGTGCAGATGGCTGATAGTTTTCACACAG

rat\_cDNA  
human 5+3 corrected  
mus\_cDNA\_5

GTCTTTACCACTGCATAAGCAACATGATGCAGATGCGGATGTTCTCATAACAGGATAA  
GCGTATATCACTGTATAAGCAGCAATTATGATGATGCAGATATTCTCACCTATAGGATAA

rat\_cDNA  
human 5+3 corrected  
mus\_cDNA\_5

CTGTGGTAGAGCCCTATAGAGAAAGCACACATGACAGTGGAGTCCAGCACACAGTGGTTA  
CTGTGGTAGAACCTTTGGTGAAGCCTATCRGGAAAAATGGGATTCTACACACAGTTTCTA

rat\_cDNA  
human 5+3 corrected  
mus\_cDNA\_5

CGGGTGAGAGCTCGAAGCTTCATGCGCTTCCAGCGGGTGTCCAGATGCTTCTATTAGCT  
TTGGTGAAACACTTGATCTTCATGCGATCTACTGGTATCCAGATGCGCTCTATTAGCT

rat\_cDNA  
human 5+3 corrected  
mus\_cDNA\_5

GGATTCTTCCAGGGAACACTGTGTTCTCTCAGCATCAAGAGACAGGCAAAATCTTAACA  
GGATTCTTCCAGGGAACAAATGTGCTCTATCAGTCATCAAGAGACAGAAAGTTCTTAACA

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

ATGGGACCTTAAGAATATTACAGGTACGCCAAAGATCAAGGTCATTACCAATGTGTGG  
ATGGCACATTAGAATATTACAGGTACCCCCGAAGACCAAGGTTATTATCGCTGTGTGG

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

CTGCCAACCCATCAGGGGCCGACTTTTTCAGTTTAAAGTTTCAGTTCAAAGAAAGGCC  
CAGCCAAOCCATCAGGGGTTGATTTTTTGATTTTCCAAGTTTCAGTCAAGATGAAAGGC

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

AAAGGATGGTTGAGCATGACAGGGAGGCAGGTGGATCTGGACTTGGAGAACCCAACTCCA  
AAAGGCCCTTGGAGCATGATGGAGAACAGAGGGATCTGGACTTGATGAGTCCAATCCTA

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

GTGTTTCOCTTAAGCAGCCAGCATCTTTGAACTCTCTGCATCAGCTTTGACAGGGTCAG  
TTGCTCATCTTAAGGAGCCACCAGGTGCACAACTCCGTACATCTGCTCTGATGGAGGCTG

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

AGGCTGGAAAACAGTCTCCGGTGTACATAGGAAGAACAAACATAGAGACTTAATACATC  
AGGTTGGAAAACACACCTCAAGCACAGTAAGAGGCACAACCTATCGGGAATTAACACTCC

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

GGCGGCGTGGGGATTCCACGCTOOGGCGATTGAGGGAGCATAGGAGGCAGCTCCCTCTCT  
AGCGACGTGGAGATTCAACACATGACGTTTTAGGGAGAATAGGAGGCATTTCCTCCCT

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

CTGCTCGGAGATTGACCGCAACGCTGGGCAGCACTTCTAGAAAAAGCCAAAAGCAATT  
CTGCTAGGAGAATTGACCCACACATTGGGGGGCACTGTTGGAGAAAGCTAAAAAGCAATG

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

CTGTGCCAAAAAGCAAGAAATACCAAGTAAGGCCAGTGCCACTGGCTGTTCCTCTCG  
CTATGCCAGACAAAGCAGAAATACCAAGTGAAGCCACCCCAAGTGGTCAOCCAACTCC

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

TGGAACTCACTGACGAGGAAAGGATGCCCTCTGGCATGATTCTCCAGATGAAGAATTCA  
CAAACATAOCTGGTGAAGAGACGATTCTCAGGCATGCTCGCTCTACATGAGGAATTTA

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

TGGTTCTGAAACTAAGGCTTCTGGTGTCCAGGAAGGTCACCAACTGCTGACTCTGGAC  
TGGTCCGGCCACTAAAGCTTTGAACCTTCAGCAAGGACAGTGAAGTCTGACTCCAGAA

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

CAGTAAATCATGGTTTTATGACGAGTAAGCTTCTGGCACAGAGTCTCAACTGTGAATC  
CAATATCTGATAGTCTATGACAAACATAAATTATGGCACAGAACTCTCGTTGTGAATT

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

CACAAACACTACAATCTGAGCACCTTCTGATTTCAATTTATTAGTGTAACAAAGGTA  
CACAAATACACCACTGAAGAACCCACAGATTTCAAACGTGTCTACTGCTATTAAACTA

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

CAGCTGTGACAAAGAGTATGAACCATCCTAGCAGCAAAATAGAAGATACACCAACC  
CAGCATGTCAAAGATATAAACCCACCATGTCAAGCCAAATACAAGGCACCAACCAATC

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

AAAACCAATCATTTATCTTTCCATC-----AGTAGCTGAAATTCGAGATTCTG-CT  
AACATTCACTCACTGTCTTTCCACTGCTACTTGGAGCAACTGAATTTCAGGACTCTGACA

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

CAGGCA---GGAAGAGCATCTTCCCAAGTGCACCCCTGTAAACAGGGGAAACATGGCT  
GAGGGAAGAGGAAGAGAGCATTTCC---AGTA-ACCCCAATAACAGTAAGGACTATGATC  
-----

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

ACCTATGGCCCTACCAACACATA--TAGTAGCTTTACCAGCAAAGCCAGTACAGTCTTGC  
AAAGATGNTCAATGTCAAANATGCTTAGTAGCACCACCAACAA-CTATTA-----TTAG  
-----

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

AGCCATTAATCCACAGAAAGTTATGGACCTCAGATACCTATTACAGGAGTCAGCAGAC  
AGTCAGTAATACCACAAATAGTCAT-----CAGACATCTGTAAGAGAAGTGAGTGAAC  
-----

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

CTAGCAGTAGTGACATCTCTTCTCAGCTACTGCAGACCCCTAGCTTCTCCAGTCACCCCT  
CCAGGCACAATCCTTCTATTTCTCAGCTACTCAAATACTTAGCACCCTCCAGCTTCCCTT  
-----

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

CAGGTTCAACACACCACTGCTCTCTTTATTTACATTCCTAGAAACAACATACAGGTA  
CAGATCCACACACAGCTGCTCATTTCTCAGTTTCOGATCCCTAGANNNAATAGTACAGTTA  
-----

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

ACTTCCCTTGTCCAGGCACTTGGGAAGAGAGAGGACAAATTTGGAGCAGAGGGAGAGTTA  
ACATCCCGCTGTTCAGACGCTTTGGGAGGCAGAGGAAATTTGGCGAAGGGGGGAGATT  
-----

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

AAAACCCACATAGAACCCAGTTCTCCGAGGCATAGACACAGGACTGTGAGGCCAGCAA  
TCAGCCCATATAGAACTCCAGTTCTCCGAGGCATAGATACAGCATTTTCAGGTCAACAA  
-----

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

TCAGGGGACCTGCTAACAAAATGTGAGCCAGTTCCAGCCACAGAGTACCCCTGGGATGT  
CCAGAGGTTCTTCTGAAGAAAGCACTACTGCTTCTCAGCCACAGTGTCTCAATGTGACAT  
-----

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

GCCACACATGTCTTCCGAGAGGGGCTCAGTGGCTACTGCAGCACTGTTCAGTTCCAA  
GTCTGTCTGTCTTCCAGGGAGAGGCTCACCCTGCCACAGCAGCATTTGTCTTTCCAA  
-----

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

GTTCTCCACAGTGGCCCTCCCAAACTAATAATGTTGGGGTCATAGCAGAAGAGTCTA  
GTGCTGTCTCCATCAGCTTCCCAAGCTGACATTGCTAGAGTCCCATCAGAAGAGTCTA  
-----

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

CCACTGTGGTCAAGAACCACTGTTACTATTTAAGGACAAACAAATGTAGATATTGAGA  
CAACTCTAGTCCAGAATCCACTATTACTACTTGAGAACAAACCCAGTGTAGANNNGAAA  
-----

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

TAATACAAACCACTACAAAATATTCGGAGGGGAAGTAACCACTGATTCCTACGGAG  
NNACAAACCCACATAAAATATTCAGGACTNGAAATTTCCCAAGTGACTCCAACTGTGT  
-----

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

CAAGCATGACTTCTGCTCCAACATCTGTATCCCTGGGGAAATCTCCTGTAGACAATAGTG  
CAGTCATGACATATGCTCCAACTCCATACCGATGGAAAAACTCACAAGTAAACGCCA  
-----

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

GTACCTGAGCCTGCTGGGACCATCCAACTGGGAAAGATTCACTGGAAACCAACAC  
GTTACCCAGTGTCTGTAGCAACCAATGAAGCTAAAGAGATTCACTGATTACATCGTCAC  
-----

rat\_cDNA  
human\_5+3 corrected  
mus\_cDNA\_5

TTCCACGCCCCCTCAGCACACCCCTCAATAOCAR-----CAAGCACAAAATTCCTAAAGA  
TTTCAGGTGCTATCACCAGGCCCAATGACTATTATAGCCATTACAAGGTTTCAAGAA

rat\_cDNA  
human\_5+3 corrected  
mus\_cDNA\_5

GGAAAACTCCCTTGACACAGATCTTGTAATAACAGAGAAGGAGGGGATGTTAAAGA  
GGAAATTCCTGGCAACAGAACTTTGTAATAACCATACCCAAAAGGCAGATTAAAGGA

rat\_cDNA  
human\_5+3 corrected  
mus\_cDNA\_5

ATCCATATCAATTGGTTTACAAAAGAACCCAGCGCGAAGGTTCCAAAATAGTCCTC  
ATCAACATAAAGTTAGTTTACAAAAGCACAGGTGTGATGCTTCTAAAACATCTCCTG

rat\_cDNA  
human\_5+3 corrected  
mus\_cDNA\_5

TTTTACCCACAG-GTCAGAGTTCCCTCAGATTCTACAACCTCTCTTGACAAGTCCSOCA  
CTTTACC-ACAGAGACAAAGTTCCCTTTCCATTTCACACACTTTCAACAGTGTGATG

rat\_cDNA  
human\_5+3 corrected  
mus\_cDNA\_5

CCAGCTCTGTCTACAACATGGCTGCCACTCAGAACAGGGCACTGAAGTAGTATCAGGT  
CAAAATCCATCTAATAOCTTGACTACCGCTCACCACACTACGACCAA--ACACACAA-T

rat\_cDNA  
human\_5+3 corrected  
mus\_cDNA\_5

GCCAGAAGTCTCTCAGCAGGGAAGAGCAGCCCTTCACCAACTCCTCTCCAG-TGCTTCC  
CCTGGAAGTCTTCCACAAAGAAGGAGCTTCCCTTC-CCACCCCTTAACCCATGCTTCC

rat\_cDNA  
human\_5+3 corrected  
mus\_cDNA\_5

TAGACOCATAAGCAAGAGATCTAATACATTAAACTTCTTGTCAACGGAAACCCCCACAGT  
TAGTATTATAAGCAAGAGACTCAGTACAAAAGCATCATATCAACGCAACAGCAACCGC

rat\_cDNA  
human\_5+3 corrected  
mus\_cDNA\_5

GACAAGTCTTACTGCTACTGCATCTGTCAATTATGTCTGAAACCCACGAACAGATCCAA  
AACAATCTTACCTTCCCTGCATCTGTCTATCATTTATGAACCCAAACAGAGAGATCTAG

rat\_cDNA  
human\_5+3 corrected  
mus\_cDNA\_5

AGAAGCAAAGACCAATAAAGGGGCTCG---GAAGAACAGAAACACGCAACACCCAC  
AGCACAAACATACAAAGAGAGAGGAOCTCAAAAGAAGAACAGGACTGACCCAAACATCTC

rat\_cDNA  
human\_5+3 corrected  
mus\_cDNA\_5

CCCCAGGCAGGTTCTGGCTATAGTGCACTACTCAGCTCTAACACAGCTGATAOCCOCTT  
TCCAGACCAAGTCTCTGGCTTCACTACACCCACTGCTATGACNACCTCCTNNGCTCTNN

rat\_cDNA  
human\_5+3 corrected  
mus\_cDNA\_5

GGCTTTCAGTCATTCCCCACGACAGATGATGGTGGAAATGTAAGTGCAGTTGCTTATCA  
NGCATTCACCTCATTCOCCACCAAGAAACACAACTGGGATTTCAAGCACAATCAGTTTCA

rat\_cDNA  
human\_5+3 corrected  
mus\_cDNA\_5

CTCAACAACCTCTCTCTGGCCA---TAACTGAAGTGTGAG-AGGTAC--AOCAGAC  
TTCAAGAAGTCTTAATCTGACAGATGTGATTGAAGAACTAGCCCAAGCAAGTACTCAGAC

rat\_cDNA  
human\_5+3 corrected  
mus\_cDNA\_5

TTTGGGAATACAAACAGCTTTGGAAACAAGTTTGTGAGCAATCACAGGAGAGTACCAC  
TTTGAAGAGCAAAATGCTTCTGAAACAACCTTTGTCCAGCAATCACACCAGAGTACCAC

rat\_cDNA  
human\_5+3 corrected  
mus\_cDNA\_5

ASTGAAGAGAGCCTCA---GACAC---ACCAACCAACTCCTCAGCAGTGGGGCGCCCCC  
AAGTAGGAAGCAATCATTAGACACTCAACCAACCACTTCTTGAGCAGCAGTGTACTCT

rat\_cDNA  
human 5+3 corrected  
mus\_cDNA\_5

AGTGCCTCACTCCTTCCCCACCTCCTTTTACTAAGGGTGIGGTTACAGACAGCAAAAGTCAC  
AATGCCAGTTCCCATCTCCCCCTCCCTTTACTCAGAGAGCAGTTACTGACAAOGTGGCGAC  
-----

rat\_cDNA  
human 5+3 corrected  
mus\_cDNA\_5

ATCAGCTTTCAGATGACGTCAAATAGAGTGGTCACCATATATGAATCTTCAAGGCACAA  
TCCCATTTCCGGGCTTATGACAAATACAGTGGTCAAGCTGCACGAATCCTCAAGGCACAA  
-----

rat\_cDNA  
human 5+3 corrected  
mus\_cDNA\_5

TACAGATCTGCAGCAACCTCAGCAGAGGCTAGCCCCAATCCTGAGATCATAACTGGAAC  
TCCNNNNNNNCA--AATGCCAAGTTCA-----CENNAATTGNGAACNNNNMACTCNNNN  
-----

rat\_cDNA  
human 5+3 corrected  
mus\_cDNA\_5

CACTGACTCTCCTCTAATCTGTTTCCATCCACTTCTGTGCCAGCACTAAGGGTAGATAA  
NACTTCATCTACNTCTAATCTGTTACATTCTACTCCCATGCCAGCACTAACACAGTTAA  
-----

rat\_cDNA  
human 5+3 corrected  
mus\_cDNA\_5

ACCACAGAATTCTAAATGGAAGCCCTCTCCTGCGCCAGAACACAAATATCAGCTCAAGTC  
ATCAGAGAATTCCAAATTAATCTCCATCTCCTGCGCCAGAAATACCAATTTGGGCACAAAC  
-----

rat\_cDNA  
human 5+3 corrected  
mus\_cDNA\_5

ATACTCCGAAACCATTTGAGAGGGGCAAAAGGCCAGCAGTAAGCATGTCCCCCCTCCTCAG  
ATACTCAGACATTGCTGAAAAGGCCAAAAGGCCAGAGTAAGCATGTTGGCT-ACTACAG  
-----

rat\_cDNA  
human 5+3 corrected  
mus\_cDNA\_5

-CCTTCCAGAGGCCAGCACTCTATGCTCACACTGGGAATACACAGAGCATGCAGAAAAGA  
GCTGTCCGAGGCCACCACTCTTGTTCAGATTGGGATGGACAGAGAACACAAAGAAGA  
-----

rat\_cDNA  
human 5+3 corrected  
mus\_cDNA\_5

GTGTTTTTGATAAGAAACCTGGTCA--AACC--CAACTTCCAAACATCTGCTTACGTCT  
GTGACTTTGATAAGAAACAGTTCAAGAGCAACAACCTCCAAACTCCTTCCCTTTGACT  
-----

rat\_cDNA  
human 5+3 corrected  
mus\_cDNA\_5

CTCTACCTAAGACTCTATTGAAAAGGCCAGAAATAATTGGAGGAAAGGCTGCAAGCTTTA  
CTTTGCTAGGTATATATTTGAAAAGGCCAGGATAGTTGGAGGAAAGCTGCAAGTTTTA  
-----

rat\_cDNA  
human 5+3 corrected  
mus\_cDNA\_5

CAGTTCAGCTAATTCAGACGTTTTCCTTCTGAGGGCTGTGGAGAACCACTGCCCA  
CTATTCCAGCTAATTCAGATGCTTCTTCTTCCCTGTGAAGCTGTGGAAATCCCTGCCCA  
-----

rat\_cDNA  
human 5+3 corrected  
mus\_cDNA\_5

TCATCCACTGGACCAAGTTTCATCAGGANTTGAATATCCRAGGGACACAGAAAGCC  
CCATTCTGGACCAAGTNNNNTCAGGACTTGATTATCTAAGAGGAAACAGAAATAGCA  
-----

rat\_cDNA  
human 5+3 corrected  
mus\_cDNA\_5

GGTTCACGTCCTTCCCAATGGCACTTGTCCATCCAGAGGGTCAGTATTCAAGACCGTG  
GGGTCCAGGTTCTCCCCAATGGTACCCCTGTCCATCCAGAGGGTGGAAATTCAGGACCGCG  
-----

rat\_cDNA  
human 5+3 corrected  
mus\_cDNA\_5

GACAGTACCTGTGCTCTGCATTTAATCCACTGGGCGTAGACCATTTTCACTCTCTTTGT  
GACAGTACTTGTGTTCCGATCCCAATCTGTTTGGCACAGACCACTTCATGTCACTTGT  
-----

rat\_cDNA  
human 5+3 corrected  
mus\_cDNA\_5

CTGTGGTTTTTTTACCGSGCAAGGATTTTGGACAGACATGTCAAGGAGATCAGSTTCACT  
CTGTGGTTTTCTATCTCTCCAGGATCCTGGAGAGAGTACCAAGAGATCAGSTTCACT  
-----

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

TTGGAACTACTGTGGAACTAAAGTGCAGAGTGGAGGGTATGCCGAGGCGCTACGGTTTCCT  
CCGGAAGCACTGTGGAACTGAAGTGCAGAGCAGAAGGTAGGCCAGCCCTACAGTTACCT

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

GGATACTTGCAAACCAAACGGTGTCTCAGAAACGGCCAGGGAAGCAGAAAGTCTGGG  
GGATTCTTGCAAACCAAACAGTTGTCTCAGAAATCATCCAGGGAGTAGGCAGGCTGTGG

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

TAAACACCTGATGGAACATTGATCATCTATAATCTGAGTCTTTATGATCGTGGTTTTTACA  
TGACGGTTGACGGAACATTGGTCTCCACAATCTCAGTATTTATGACCGTGGCTTTTACA

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

AGTGTGTGGCCAGCAACCCATCTGCCAGGATTCACTGTGGTTAAGATACAGTCATCA  
ATGTGTGTGGCCAGCAACCCAGGTGGCCAGGATTCACTGCTGTTAAATACAGTCATTG

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

CAGCTCCOCTGTCTATTATAGACCAAAGAGGCAAGCCATCGTTGGGGTTTTAGTGTGAA  
CAGCACCACCTGTATTCTAGAGCAAAGGAGGCAAGTCATTGTAGGCACCTTGGGGTGA

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

GTTTGAACTGCCCTGCACTGCAAAGGAATCCCGAGCTAGTGTCTCACTGGGTCTCTT  
GTTTAAACTGCCCTGTACTGCAAAGGAATCCCTCAGCCAGCGTTTACTGGGTCTCT

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

ATGATGGGACTGAACTAAACCATTCAGTTGACTCATTCCAGATTTTCTGTATCCAA  
CTGATGGCACTGAAGTGAAACCATTCAGTTTACCAATCCAGTTGTTCTTATTTCAA

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

ATGGAACTCTGTATATAAGAACATCGCTCCTTCAGTGAGGGGCACCTATGAGTGCATTG  
ATGGGACTTTGTATATAAGAACCTAGCCTCTTCAGACAGGGGCACCTATGAATGCATTG

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

CCACAGCTCCTCAGGCTCAGAGAGAGGGTAGTGATTCTTACTGTGGAAGAGGGAGAGA  
CTACAGTTCCACTGGTTCGGAGGAGAGAGTAGTAATGCTTACAATGGAAGAGCGAGTGA

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

CAATCCOCCAGGATAGAACTGCTCTCAGAAATGGACTGAGGTGAATTTGGGTGAGAAAT  
CCAGCCOCCAGGATAGAACTGCATCCAGAAAGGAGTGAAGTGAATTTTGGGGACAAAT

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

TACTACTGAACTGCTCAGCTACTGGGGATCCAAAGCCTAGAATAATCTGGAGGCTGCCAT  
TACTACTGAACTGCTCAGCCACTGGGGAGCCCAACCCCAATAATGTGGAGGTTACCAT

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

CCAGGGCTGTCTCGACCACTGGCACAAGATGGGCGCCGAATCCAGCTACCCCAATG  
CCAGGGCTGTCTCGACCACTGG-----GCAGCTGGATCCAGCTACCCCAATG

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

GATCCTTGGTGGTTGGGTCACTGACGGAAAAAGACGCTGGTGACTACTTATGTGTGGCAA  
GATCCCTGTTTATTGGATCAGTACAGAAAAAGACAGTGGTGTCTACTTGTGTGTGGCAA

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

GAAACAAATGGGAGATGACCTAGTCTGTATCATGTCCGCTGAGATTGACACCTGCCA  
GAAACAAATGGGGATGATCTGATACATGATCATGTAGCCTAAGACTGAAACCTGCCA

rat_cDNA human 5+3 corrected mus_cDNA_5	AAATTGAACAGAGCAGTATTTTAAGAGCAAGTGCTCCATGGGAAAGATTTCAGTGT AAATTGACCAACAGCAGTATTTTAGAAGCAAGTGCTCCATGGGAAAGATTTCAGTAG -----
rat_cDNA human 5+3 corrected mus_cDNA_5	ACTGCAAGGCTCTGGCTCCCTGTGCTGAGGTATCCTGGAGTTTGCTGATGGGACAG ATTGCAAGCTTCCGGCTCCOCAGTGCCAGAGATATCTTGGAGTTTGCTGATGGAACA -----
rat_cDNA human 5+3 corrected mus_cDNA_5	TGCTCAACAATGTAGCCCAAGCTGATGACAGTGGCTATAGGACCAAGAGGTACACCTTT TGATCAACAATGCAATGCAAGCCGATGACAGTGGCCACAGGACTAGGAGATATACCTTT -----
rat_cDNA human 5+3 corrected mus_cDNA_5	TCCACAATGGAACTTGTATTTCAACAAGCTTGGGATGGCAGAGGAAGGAGATTATCT TCAACAATGGAACTTTATACTTCAACAAGTTGGGGTAGCGGAGGAAGGAGATTATACTT -----
rat_cDNA human 5+3 corrected mus_cDNA_5	GCTCTGCCAGAACACCTTAGGGAAAGATGAGATGAAGTCCACCTAACAGTTCTAACAG GCTATGCCAGAACACCTTAGGGAAAGATGAATGAAGGTCCACTAACAGTTATAACAG -----
rat_cDNA human 5+3 corrected mus_cDNA_5	CCATCCACGGATAAGGCRAAGCTACAGACCACCATGAGGCTCAGGGCTGGAGAACAG CTGCTCCCCGATAAGGCAGAGTAACRAAACCAACAGAGAATCAAAGCTGGHACACAG -----
rat_cDNA human 5+3 corrected mus_cDNA_5	CTGTCTTGACTGCGAGTCACTGGGGAACCGAAGCCCAATGTATTTTGGTTGCTGCCTT CTGTCTTGACTGTGAGGTCACTGGGGATCCCAACCAAAAATATTTTGGTTGCTGCCTT -----
rat_cDNA human 5+3 corrected mus_cDNA_5	CCAACAATGTCATTTTCTTCTCAATGACAGGTTCACATTTCTGCAATAGACTTTGT CCATGACATGATTTCTTCTCTATTGATAGGTACACATTTCTGCAATGGGTCTTTGA -----
rat_cDNA human 5+3 corrected mus_cDNA_5	CCATCCATAAAGTGAAACCACTTGACTCTGGGACTATGTGTGCTAGCTCAGATCCTA CCATCAACAAGTGAAACTGCTCGATTCTGGAGAGTACGTATGTGTAGCCCGAAATCCCA -----
rat_cDNA human 5+3 corrected mus_cDNA_5	GTGGGGATGACACTAAGACATACAACTGGACATTGTCTTAACCTCCATTAACTAATG GTGGGGATGACACCAAAATGTACAACTGGATGTGTCTCTAAACCTCCATTAACTAATG -----
rat_cDNA human 5+3 corrected mus_cDNA_5	GCCTGTATGCAACAAGACTGTATTAAAGCCACAGCCATTGGGCACTCCAAAAATACT GTCTGTATACAAACAGAACTGTATTAAAGCCACAGCTGTGAGACATTCCAAAAACACT -----
rat_cDNA human 5+3 corrected mus_cDNA_5	TTGACTGCAGAGCAGATGGGATCCCATCTTCCAGGTCACTGGATTATGCCAGGCAATA TTGACTGCAGAGCTGAAGGGACACCATCTCTGAGTCATGTGATCATGCCAGCAATA -----
rat_cDNA human 5+3 corrected mus_cDNA_5	TTTTCTCCAGCTCCATACTTTGGAAGCAGAGTCAGGTCCATCCAAATGGAACTTGG TTTTCTCAGGCCCCATACTATGGAAGCAGAATCAGTCCATAAAATGGAACTTGG -----
rat_cDNA human 5+3 corrected mus_cDNA_5	AGATGAGGAACATCCGGCTTTCTGACTCTGCGGACTTCACCTGTGTGGTTCGGAGGAGG AAATTAGGAATGTGAGGCTTTCAGATTGAGCGACTTTATCTGTGTGGCCCGAAATGAAG -----



GAGGAGAGAGTGTGTGGTAGTGCAGITAGAAGTCCTAGAAATGCTGAGAGAGACCAACAT  
GTGGAGAGAGACGTGTGGTAGTACGTTAGAGAGTACTGGAAATGCTGAGAGAGACCGACAT

TCAGAAACCCATTCAACGAAAAAGTCATCGCCCAAGCTGGCAAGCCCGTAGCACTGAACT  
TTAGAAATCCATTTAATGAAAAATAGTTGCCACAGCTGGGAAAGTCCACAGCAATTGAATT

GCTCTGTGATGGGAACCCCACTGAAATTACCTGGATCTTACCTGAOGGCACACAGT  
GCTCTGTGTGATGGTAAACCAACCACTGAAATAATCTGGATTTTACCAAAATGGCACACGAT

TTGCTAACAGACACACAATTCCCGTATCTGATGGCAGGCAATGGCTCTCTCATCCTTT  
TTTCCAATGGACCACAAAGTTATCAGTATCTGATAGCAAGCAATGGTCTCTTTATCATTT

ACAAAGCAACTCGGAACAGTCAAGGCAAGTATCGCTGTGCAGCCAGGAATAAGGTTGGCT  
CTAAACAACTCGGAGGATGCAGGAAAATATCGCTGTGCAGCTAGGAATAAGTGGCT

ACATCGAGAACTCATCCTGTTAGAGATTGGGCAGAAAGCCAGTCATTCTGACATACGAAC  
ATATTGAGAAATTAGTCATATTAGAAATTGGCCAGAAAGCCAGTTATTCTTACCATATGCAC

CAGGGATGGTGAAGAGCGTCAGTGGGGAAACCGTTATCACTGCATTGTGTCTCTGATGGGA  
CAGGGACAGTAAAGGCATCAGTGGAGAATCTCTATCACTGCATTGTGTCTCTGATGGAA

TCCCCAAGOCAATGTCAAGTGGACTACACCGGGTGGCCATGTAATCGACAGGCOCTCAG  
TCCCTAAGCCAAATATCAATGGACTATGCCAAGTGGTTATGTAGTAGACAGGCOCTCAA

TGGATGGAAATACATACTGCATGAAAATGGCAGCTGGTCATCAAGCAACAACAGCTC  
TTAATGGGAAATACATATTGCATGACAATGGCACTTAGTCATTAAAGAGCAACAGCTT

ACGACCAAGGAATTATATCTGTAGGGCTCAAAACAGTGTGGGCAGGCAGTTATTAGCG  
ATGACAGAGGCAACTATATCTGTAAAGGCTCAAAATAGTGTGGTCAATACACTGATTACTG

TGTCAGTGAATGGTTGTGGCCTACCGTCCCGAATCATAAACTACCTACCCAGGAACATGC  
TTCCAGTAATGATGTAGCCTACCGTCCCGAATTACBBATGTTCCACCCGGAGTATTG

TCAGGAGGACAGGGGAGCCATGCAGCTCCACTGTGTGGCCTTGGGAATCCCAAGCCAA  
TCACCAGGACAGGGGACGCCCTTTGAGCTCCACTGTGTGGCCTTGGGAGTTCCCAAGCCAG

AAGTCACTGGGAGACGGCAAGACACTCCCTGCTCTCAAAGCAACAGCAAGTAAACCCC  
AAATCACAATGGGAGATGCGTGACCACTCCGTTCTCTCAACGGTAAGTAAAGACAGGACAC

ATAGAAGTGAGATGCTTCACCCACAGGTACGCTGGTCATTGAGAACTCTCCAAACCTCGG  
ATGGAAGTGAGCAGCTTCACCTACAGGTACCTAGTCATTGAGAACTCCCAACCTCGG

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

ATTCCGGAGTCTATAAGTGCAGAGCTCAGAACCTACTTGGGACTGATTACGCCAACACTT  
ATTCTGGGATATACAAATGCACAGCAAAGAACCCACTTGGTAGTGATTATGCAGCAACGT

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

ACATCCAGGTACTCTGACAGGAAGGGGGAGACTAAAATTCAACAGAAAGTCCACATCCACA  
ATATTCAAGTAATCTGACATGAA----ATAATAAAGT-CAACAA----CATCTGGGCA

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

GGGTTTATTTTGGAGAAGTTTAAATCAAAGGCAGCCATAGGCATGTAAATGAGTCTGA  
GAATTTATTTTGGAGAAGTTTAAATCAAAGGCAGCCATAGGCATGTAAATGAATTTGA

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

ATACATTACAGTATTAAATTTACAATGGACATGCCA--TGA--GACTTGTAATGAAA  
ATACATTACAGTATTAAATTTACAATGAACATGCRAAATAAAAGGACTTGTAATAAAT

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

GCATTGTGAACAGAA--ACCGAGTCTCTG--TGGATCTCAAAGCRAACTCTTAACCTAA  
GCATTATGAAGTATGATGACTGATTATTTAATGGATCTCAAACCAAACCTTTTAACTTA

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

GGCAGTTTGATTTTGCCAAACAATAAACAACATTAAGAGAAAAAATGATCCACTAC  
GGCAGTTTATTTTGCCAAACAATAACAATAAACAA--ACATTGAACGGTTCACAT

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

GAAATAACAAACGGCTAATGCACCTGAATTCT-CAGTAAAAAGACCTTTCTCTGGCTAAC  
AAAATAACAAATGGCTAATGTACCTGAATTTTTCAGTAAAAAA--TGAACCTT-CTAAT

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

AGTTGCCAGCTGCCTCGTCTCTGTTTCCTACCAATGTCAAAACATCGCACACAGGCTGA  
A--CCAGTTGCCTAGTGTCCACCTCTCTATCAATGTTACAAGCATGGCACTCAG--A

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

ATGGAGTCAACGGGAAGATTAAAGTTTCCGGTCTGTGTAATCTCAATGTACAAATATTC  
ACAGAGACAAATGGAAATATTAAATCTGC-----AATCT--ATGTATAAATATT

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

TGTCTCTGGTTTATAAACATTTT-GATAAAACCGAAAAA-----  
TGT----GGTTTATAAATTTTCTGCTAAACCTACAGAAATAG

rat\_cDNA  
human\_5+3\_corrected  
mus\_cDNA\_5

AAAA  
-----  
-----

(rat\_cDNA: SEQ ID NO:7)

(human\_5+3 corrected: SEQ ID NO:8)

(mus\_cDNA\_5: SEQ ID NO:9)

Figure 15

rat	MQVRGREVSGLLISLTAVCLVVPFSGRACPRRCACYPTEVHCTFRYLTSIPDGIPANVE
human_5+3_corrected	MKVKGRTITCLLVSEFAVICLVATPGGKACPRRCACYMPTEVHCTFRYLTSIPDSIPPNVE
mouse_5_corrected	MQKRGREVSCLLISLTAVCLVVPFSGRVCPRRCACYPTEVHCTFRDLTSIPD-GPANVE
rat	RINLGYNLSLRLTENDFDGLSKLELIMLHSGIHRVSDKTFSGLQSLQVLKMSYNKVQII
human_5+3_corrected	RINLGYNLSLRLMETDFSGLTKELEIMLHSGIHTIPDKTFSDLQALQVLKMSYNKVRKL
mouse_5_corrected	RVNLGYNLSLRLTENDFDGLSKLELIMLHSGIHRVSDKTFSGLQSLQVLKMSYNKVQII
rat	RKDTFYGLGSLVRLHLDENNIEFINPEAFYGLTSLRLVHLEGNRLTKLHPDTFVSLSYLQ
human_5+3_corrected	QKDTFYGLRSLTRLHMDENNIEFINPEVYGLNRLVHLEGNQLTKLHPDTFVSLSYLQ
mouse_5_corrected	EKDTLYGLRSLTRLHLDENNIEFINPEAFYGLTLRLVHLEGNRLTKLHPDTFVSLSYLQ
rat	IFKTSFIKYLELSDNFLTSLPKEMVSYMPNLESYLHGNPWTCDCHLKWLSEWMQGNPDI
human_5+3_corrected	IFKISFIKELYLSDNFLTSLPKQEMVSYMPDLDSLYLHGNPWTCDCHLKWLSDWIQ--PDV
mouse_5_corrected	IFKTSFIKLYLYDNF-TSLPKEMVSSMPNLESYLHGNPWTCDCHLKWLSEWMQGNP--
rat	IKCKKDRSSSSPQQCPLCMNPRI SKGRPFAMVPSGAFLCTKPTIDPSLKSLSVTOEDNG
human_5+3_corrected	IKCKKDRSPSSAQQCPLCMNPRTSKGKPLAMVSAAFQCAKPTIDSSLSKSLTILEDSS
mouse_5_corrected	-----
rat	SASTSPQDFIEPFGSLSLNMTKXSGNKADMVCSIQKPSRTSPATTEENDYIMLNASFST
human_5+3_corrected	SAFISPGQGMAPPFGSLTAMTDQSGNEANMVCSIQKPSRTSPATTEENDYIVLNTSFST
mouse_5_corrected	-----
rat	NLVCSDVDYNIQPVKQLLALYSDSPILILKPKQLTETPSLSSRYKQVALRPEDIFTSIEA
human_5+3_corrected	FLVCNIDYGHQIPVWQILALYSDSPILILERSHLLSETPOLYYKYKQVAPKPEDIFTNIEA
mouse_5_corrected	-----
rat	DVRADFFWFQQEKIVLQNLRTATTLSTLQIQFSTDAQIALPRAEMRAERLKWTMILAMNN
human_5+3_corrected	DLRADPSWLMQDQISLQNLRTATTLSTLQIQYSSDAQITLPRAEMLRPVKHKWTMISRDNN
mouse_5_corrected	-----
rat	PKLERTVLVGGTIALSCPGKGDPSPHLENLLADGSKVRAPYVSEDGRILLIDKNGKLELQ
human_5+3_corrected	TKLEHTVLVGGTVGLNCPGQGDPTPHVDLLADGSKVRAPYVSEDGRILLIDKSGKLELQ
mouse_5_corrected	-----
rat	ADSFDAGLYHCISTNDADADVLTIRITVVEPYGESTHDSGVQHTVVTGETLDLPCLSTGV
human_5+3_corrected	ADSFDTGVYHCISSNYDDADILTIRITVVEPLVEAYQENGIEHTVFIGETLDLPCHSTGI
mouse_5_corrected	-----
rat	PDAISWILPGNTVFSQPSRDRQILNNGTLRLILQVTPKDOGHYQCVAAANPSGADFSSFKV
human_5+3_corrected	PDAISWVIPGMNVLYOSSRDKKVLNNGTLRLILQVTPKDOGYRQVAAANPSGVDFLIFQV
mouse_5_corrected	-----
rat	SVQKKQORMVEHDEAGGSGLGEPMSSVSLKQPASIKLSASALTGSEAGROVSGVHRNK
human_5+3_corrected	SVKMKQORPLEHDGETEGSGLDESNPIAHLEKPPGAQLRTSALMEAEVGKHTSSTSRNK
mouse_5_corrected	-----
rat	HRDLIHRRRGDSLTRRFREHRRQLPLSARRIDPQWAALEKAKKNSVPKQENTTVKPV
human_5+3_corrected	YRELTLQRRGDSLTRRFRENRRHFPSSARRIDPQWAALEKAKKQAMPDKRENTTVSP
mouse_5_corrected	-----



rat -TSKHLPPYVSLPKTLKKPRIIGGKAASFTVPANSDVFLPCEAVGDPLPIIHMTRVSSGX  
human\_5+3\_corrected TTskLLPFDLSRYIFEXPRIVGGKAASFTIPANSDAFLPCEAVGNPLPTIHWTRVS-GL  
mouse\_5\_corrected -----

rat EISQGTQKSRPHVLPNGTLSIQRVSIQDRGQYLCSAFNPLGVDHFFHVLSVVFYFARILD  
human\_5+3\_corrected DLSRGNQNSRVQVLPNGTLSIQRVSIQDRGQYLCSASNLFGTDHLHVLSVVSYPRIIE  
mouse\_5\_corrected -----

rat RHVKRITVHFGSTVELKCRVSGMPRPTVSWILANQTVVSETAKGSRKVVVTPDGTLLIYN  
human\_5+3\_corrected RRTKEITVHSGSTVELKCRAEGRPSPTVTWILANQTVVSESSQGSROAVTVVDGTLVLHN  
mouse\_5\_corrected -----

rat LSLYDRGFYKCVASNPGGQDSLIVKIQVITAPPVITEQKROAIVGLGGSCLKLPCTAKGT  
human\_5+3\_corrected LSIYDRGFYKCVASNPGGQDSLIVKIQVIAAPPVILEQRRQVIVGTWGESCLKLPCTAKGT  
mouse\_5\_corrected -----

rat PQPSVHWVLYDGTTELKPLQLTHSRFFLYPMGTLYIRSIAPSVRGTYECIATSSSGSERRV  
human\_5+3\_corrected PQPSVYVWVLSDGTVEVKPLQFTNSKLFLFSNGTLYIRNLASSDRGTYECIATSSSGSERRV  
mouse\_5\_corrected -----

rat VILTVEEGETIPRIETASQKNTVNLGEKLLNCSATGDPKPRIIWRPLPSKAVIDQWHRM  
human\_5+3\_corrected VMLTMEERVTSPRIEASQKNTVNLGEKLLNCSATGEPKPOIMWRPLPSKAVVDQ----  
mouse\_5\_corrected -----

rat GSRIHVYPNGSLVVGSVTEKDAGDYLCVARNKMGDDLILMHVRLRLTPAKIEQKQYTKKQ  
human\_5+3\_corrected GSWIHVYPNGSLFVTEKDSGVYLCVARNKMGDDLILMHVSLRLKPAKIDHKQYTKKQ  
mouse\_5\_corrected -----

rat VILGKDFQVDCRAGSGSPVPEVSNLPGDGLVNNVAQADDSGYRTKRYTLEHNGTLYFNVY  
human\_5+3\_corrected VILGKDFQVDCRAGSGSPPEISWSLPGDTMINNVAQADDSGHRTRRYTLEFNNGTLYFNKY  
mouse\_5\_corrected -----

rat GMAEEGDYTCQAQNTLGKDEMKVHLTVLTAIPRIQSYRTTMRIRAGETAVIDCEVTIGEP  
human\_5+3\_corrected GVAEEGDYTCQAQNTLGKDEMKVHLTVITAAPRIQSNKTKRIKAGDTAVLDCEVTIGDP  
mouse\_5\_corrected -----

rat KPNVFWLLPSNNVISFSNDRFTTHANRLSIHKVKPLDSGDYVCVAQNPSGDDTKYTKLD  
human\_5+3\_corrected KPKIFWLLPSNDMISFSIDRYTFHANGSLTINKVKLLDSGEYVCVARNPSGDDTKMYTKLD  
mouse\_5\_corrected -----

rat IVSKPEPLINGLYANKTVIKATAIHSKKYFDCRADGIPSSQVWIMPGNIFLPAFYGSR  
human\_5+3\_corrected VVSKEPLINGLYTNRTVIKATAVRSKKHFDCAEGTPSPEVMWIMPDNIFLTAIFYGSR  
mouse\_5\_corrected -----

rat VTVHPNGTLEMRNIRLSDSADFTCVVRSEGGESVLVVQLEVLEMLRRPTFRNPFNEKIVA  
human\_5+3\_corrected ITVHKNGTLEIRNIRLSDSADFTCVARNEGGESVLVVQLEVLEMLRRPTFRNPFNEKIVA  
mouse\_5\_corrected -----

rat QAGKPVAINCSVDGNPPPEITWILPDGTQFANRPHNSPYLMAGNSLILYKATFNKSGKY  
human\_5+3\_corrected QLGKSTALNCSVDGNPPPEIILWILPNGTRFSNGPQSYQYLIASNGSPFIISKTTREDAGKY  
mouse\_5\_corrected -----

rat RCAARNKVGYIEKLIILEIGQKPVILTYEPGMVKSVSCEPLSLHCVS DGIPKPNVKTTP  
human\_5+3\_corrected RCAARNKVGYIEKLVILEIGQKPVILTYAPGTVKGISGESLSLHCVS DGIPKPNIKWTMP  
mouse\_5\_corrected -----

rat GGHVIDRPQVDGKYILHENGTLVIKATTAHDQGYICRAQNSVGQAVISVSVMVAITRK  
human\_5+3\_corrected SGYVVDRPQINGKYILHDNGTLVIKEATAYDRGNYICKAQNSVGHTLITVPVMIVATPPR  
mouse\_5\_corrected -----

rat IINYLEFNMLRRTGEANQLHCVALGIPKPKVTWETPRESLLSKATARKPHRSEMLHPQGT  
human\_5+3\_corrected ITNRPPRSIVTRTGAAPQLHCVALGVFKPEITWEMPDSLLSTASKERTHGSEQLHLQGT  
mouse\_5\_corrected -----

rat LVIQNLQTSDSGVYKCRAQNLGTDYATTYIQV  
human\_5+3\_corrected LVIQNPQTSDSGIYKCTAKNPIGSDYATTYIQV  
mouse\_5\_corrected -----

(rat: SEQ ID NO:10)

(human\_5+3\_corrected: SEQ ID NO:11)

(mouse\_5\_corrected: SEQ ID NO:12)



[illegible]



rat	SRIHVYPNGSLVVGSVTERKDGAYLCVARNKMGDOLVIMHVRLRLTPAKIEQKQYFKKQV
human_5+3_corrected	SWIHVYPNGSLFLGSVTERKDSGYLCVARNKMGDOLVIMHVSLRLPKAKIDHKQYFRKQV * *****;*****;* *****;**** *.****;*****;***
rat	LHGKDFQVDCASGSPVPEVSWSLPDGTVLNNVAQADDSGYTKRYTLFNGTLYFNNG
human_5+3_corrected	LHGKDFQVDCASGSPVPEISWSLPGDTMINNAMQADDSGHETRYTLFNGTLYFNKVG *****;*****;***. *****;***;*****;*****;***
rat	MAEEGDYICSAQNTLGKDEMKVHLTVLTAIPRIQSYKITHRLPAGETAVIDCEVTGEPK
human_5+3_corrected	VAEEGDYTCYAQNTLGKDEMKVHLTVITAAPRIQSNKTKRIKAGDTAVLDCEVTGDPK ;***** * *****;*** ***** **. *;***;*****;***
rat	PNVFWLLPSMNVISFSNDRFTFHANRTLSIHKVKPLDSGDYVCVAQNPSGDDTKTYKLDI
human_5+3_corrected	PKIFWLLPSNDMISFSIDRYTFHANGSLTFINKVLLDSGEYVCVARNPSSGDDTKMYKLDV *;;*****;**** **;***** ;*;*;*** *****;***** *****
rat	VSKPPLINGLYANKTVIKATAIRBSKKYFDCRADGIPSSQVWIMPONIPLPAPYFGSRV
human_5+3_corrected	VSKPPLINGLYTNRTVIKATAVRHSKKHFDCAEGTPSPVEMWIMPDNIFLTAPYFGSRI *****;***;*****;*****;*****;* **;* *****;*****;***
rat	TVHPNGTLEMRNIRLSDSADFTCVVRSEGGESVLVVQLEVLEMLRRPTFRNPFNEKIVAQ
human_5+3_corrected	TVKNGTLEIRNVRLSDSADFTCVARNEGGESVLVVQLEVLEMLRRPTFRNPFNEKIVAQ *** ****;***;***** **;*****;*****;*****;*****;***
rat	AGKPVALNCSVDGNPPPEITWLLPDGTQFANRPHNSYFLMAGNGLILYKATRNKSGKYR
human_5+3_corrected	LKSTALNCSVDGNPPPEIWLNPGRFSNGPQSYQLIASNGSFIISKATTREDAGKYR **;***** *****;***;* *; **;*;***;*; *;***;****
rat	CAARNKVGYIEKLILLEIGQKPVILTYEPMVKSVSCEPLSLHCVS DGI PKPNVKWTTPG
human_5+3_corrected	CAARNKVGYIEKLVIILEIGQKPVILTYAEGTVKGISGESLSLHCVS DGI PKPNIKWTMP *****;***** ** **;*** *****;*** *
rat	GEVIDRPQVDGKYILHENGTLVIKATTAHQGNYYICRAQNSVGQAVLBSVHVVAYPRI
human_5+3_corrected	GYVVD RPQINGKYILLHDNGTLVIKATAYDRGNYYICRAQNSVGHTLITVPVMIVAYPRI *;*;*****;*****;***** **;*;*****;*****;***;*;*****
rat	INYLFRNMLRRTGEAMQLHCVALGIPKPKYTWETPRHSLLSKATARKPHERSEMHPQGTIL
human_5+3_corrected	TNRPPRSIVTRTGAAPQLHCVALGVKPEITWEMPDHSLSTASKERTHGSEQLHIGTIL * **;.. ** *;*****;***;*** * *****;* ..* ** ** *
rat	VIONLQTSDSGVYKCAQNLLGTDYATYIQVL
human_5+3_corrected	VIONPQTSDSGIYKCTAKNPLGSDYATYIQVI *** *****;*** *;***;*****;*****;

(rat: SEQ ID NO:13)

(human\_5+3\_corrected: SEQ ID NO:14)

Figure 17

MQKRGREVSCLLISLTAICLVVTPGSRVCPRRCACYPTEVHCTFRDLTSIPDGPANVER  
VNLGYNSLTRLTENDFSGLSRLELLMLHSNGIHRVSDKTFSGLSLQVLKMSYNKVQIE  
KDTLYGLRSLTRLHLDHNNIEFINPEAFYGLTLLRLVHLEGNRLTKLHPDFTVSLSYLQIF  
KTSFIKXLYLYDNFTSLPKEMVSSMPNLESLYLHGNPWTCDCHLKWLSEWMQGNP  
(SEQ ID NO: 15)

Figure 18

MKVKGRGITCLLVSFAVICLVATPGGKACPRRCACYPTEVHCTFRYLTSPDSIPPNVE  
RINLGYNSLVRLMETDFSGLTKELELLMLHSNGIHTIPDKTFSDLQALQVLKMSYNKVRK  
LQKDTFYGLRSLTRLHMDHNNIEFINPEVFYGLNFLRLVHLEGNQLTKLHPDFTVSLSYL  
QIFKISFIKFLYLSDNFLTSLPQEMSYPDLDSLLYLHGNPWTCDCHLRWLSDWIQPDVI  
KCKKDRSPSSAQQCPLCMNPRTSKGKPLAMVSAAAFQCAKPTIDSSLKSKSLTILEDSSS  
AFISPQGFMAPFGSLTLNMTDQSGNEANMVCISIQPSRTSPIAFTEENDYIVLNTSFSTFL  
VCNIDYGHIQPVWQILALYSDSPLILERSHLLSETPQLYYKYKQVAPKPEDIFTNIEADLR  
ADPSWLMQDQISLQLNRTATTFSTLQIQYSSDAQITLPRAEMRPVKHKWTMISRDNNTK  
LEHTVLVGGTVGLNCPGQGDPTPHVDWLLADGSKVRAPYVSEDGRILIDKSGKLELQM  
ADSFDTGVYHCISSNYDDADILTYRITVVEPLVEAYQENGIHHTVFIGETLDLPCHSTGIP  
DASISWVIPGNNVLYQSSRDKKVLNNGTLRILQVTPKDQGYRCVAANPSGVDFLIFQV  
SVKMKGQRPLEHDGETEGSGLDESNPJIAHLKEPPGAQLRTSALMEAEVKGHTSSTSKRH  
NYRELTQLRRGDSTHRRFRENRHFPSPARRIDPQHWAALLEKAKKNAMPDKRENTTV  
SPPPVTQLPNIPGEEDDSSGMLALHEEFMVPATKALNLPARTVTADSRTISDSPMTNIN  
YGTEFSPVVNSQILPPEPTDFKLSTAILTTAMSKNINPTMSSQIQGTTNQHSSTVFPLLLG  
ATEFQDSDQMGRGREHFQSRPPITVRTMIKDVNVKMLSSTTNKLLLESVNSHQTSVREV  
SEPRHNHFYSHTTQILSTSTFSPDPHTAAHSQFPIPRNSTVNIPLFRRFGRQRKIGGRGRIIS  
PYRTPVLRHRHYSIFRSTTRGSSEKSTTAFSATVLNVTCLSCLPRELTTATAALSFPSAAPI  
TFPKADIARVPSEESTTLVQNPLLLLENKPSVEKTTPTIKYFRTEISQVTPTGAVMTYAPT  
SIPMEKTHKVNASYPRVSSSTNEAKRDSVITSSLSGAIKPPMTIIAITRFSRRKIPWQQNFV  
NNHNPKGRLRNQHKVSLQKSTAVMLPKTSPALPQRQSSPFHFRRLSTSVMQIPSNLTLT  
AHHTTTKTHNPGSLPTKKELPFPLNPMLPSIISKDSSTKSIISTQTAIPATTPTFPASVITYE  
TQTERSRAQTIQREQEPQKKNRDTPNISPDQSSGFTTPTAMTPPALAFTHSPPENTTGISST  
ISFHSRTLNLTDVIEELAQASTQTLKSTIASSETLSSKSHQSTTTRKASLDTPIPPFLSSSAT  
LMPVPISPFFTQRAVTDTRGDSHFRLMTNTVVKLHESSRHNLMQMPSSQLEPLTSSTSNLL  
HSTPMPALTTVKSQNSKLTSPSPWAEQFWHKPYSDIAEKGKKPEVSMLATTGLSEATTLV  
SDWDGQKNKKSDFDKPKVQEATTSKLLPFDLSRYIFEKPRIVGGKAASFTIPANSDAF  
LPCEAVGNPLPTIHWTRVSGLDLSRGNQNSRVQVLPNGTLSIQRVEIQDRFYLCASNL  
FGTDHLHVTLSVVSYPPIRERRTKEITVHSGSTVELKCRAEGRPSPTVTWILANQTVVSE  
SSQGSRQAVVTVDGTLVLHNLSTYDRGFYKCVASNPGGQDSLLVKIQVIAAPPVILEQRR  
QVIVGTWGESLKLPTAKGTPQPSVYVWVLSDGTEVKPLQFTNSKLFLFSNGTLYIRNLAS

SDRFTYECIATSSTGSERRVVMLTMEERVTSPIREAASQKRTEVNFGDKLLLNC SATGEP  
 KPQIMRLPSKAVVDQGSWIHYPNGSLFIGVTEKDSGVYLCVARNKMGGDDLILMHVSLR  
 LKPAKIDHKQYFRKQVLHGKDFQVDCKASGSPVPEISWSLPDGTMINNAMQADDSGHR  
 TRRYRLFNNGTLYFNKVGVAEEGDYTCYAQNTLGKDEMKVHLTVITAAPRIRQSNKTN  
 KRIKAGDTAAVLDCEVTGDPKPKIFWLLPSNDMISFSIDRYTFHANGSLTINKVKLLDSG  
 EYVCVARNP SGDDTKMYKLDVVSKPPLINGLYTNRTVIKATAVRHSHKHFD CRAEGTP  
 SPEVMWIMPDNIFLTAPYYGSRITVHKNGTLEIRNVRLSADFICVARNEGGE SVLVVQLE  
 VLEMLRRPTFRNPRNPFNEKIVAQLGKSTALNCVSVDGNPPPEI IWLPNGTRFSNGPQSY  
 QYLIASNGSFIISKTTREDAGKYRCAARNKVG YIEKL VILEIGQKPVILTYAPGTVKGISGE  
 SLSLHC VSDGIPKPNIKWTMPSGYVVD RPQINGKYILHDNGTLVIKEATA YDRGN YICKA  
 QNSVGHTLITVPVMIVA YPPRITNRPPRSIVTRTGAA FQLHCVALGV PKPEITWEMPDHS  
 LLSTASKERTHGSEQLHLQGT LVIQNPQTSDSGIYKCTAKNPLGSDYAATYIQVI (SEQ ID  
 NO: 16)

Figure 19

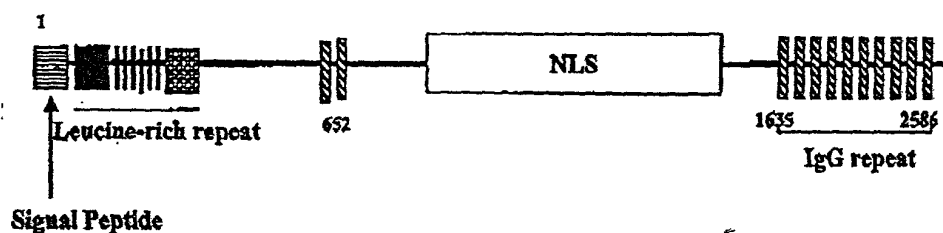


Figure 20

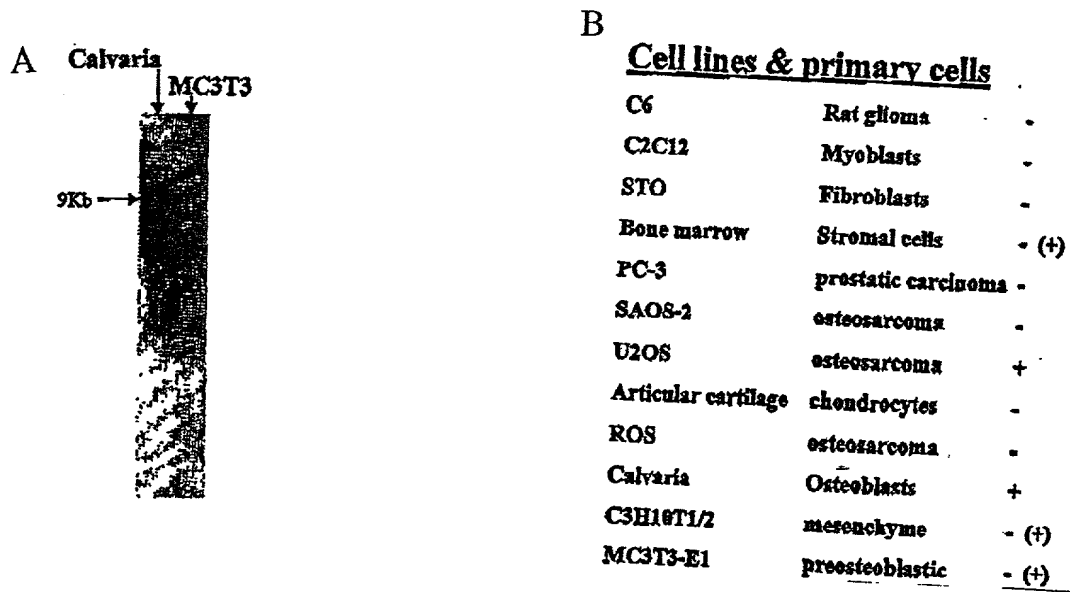


Figure 21

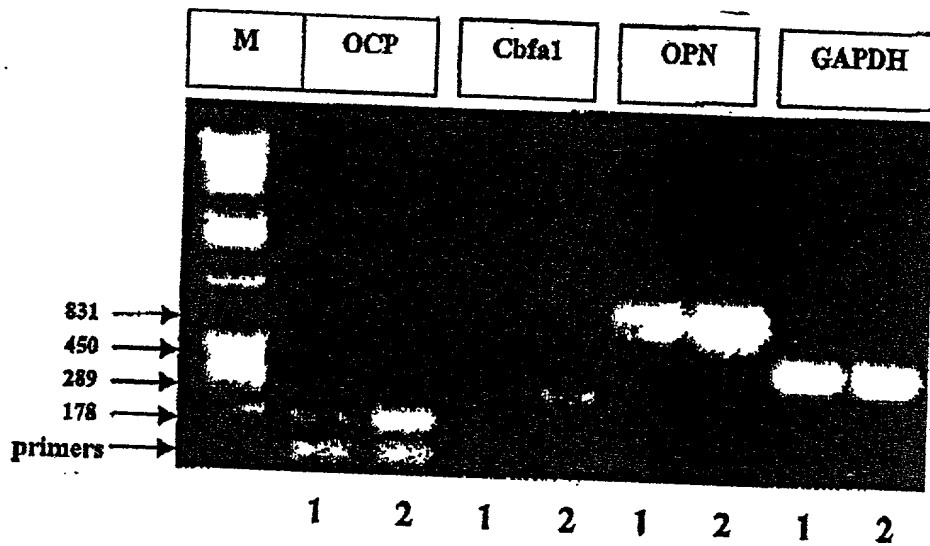


Figure 22

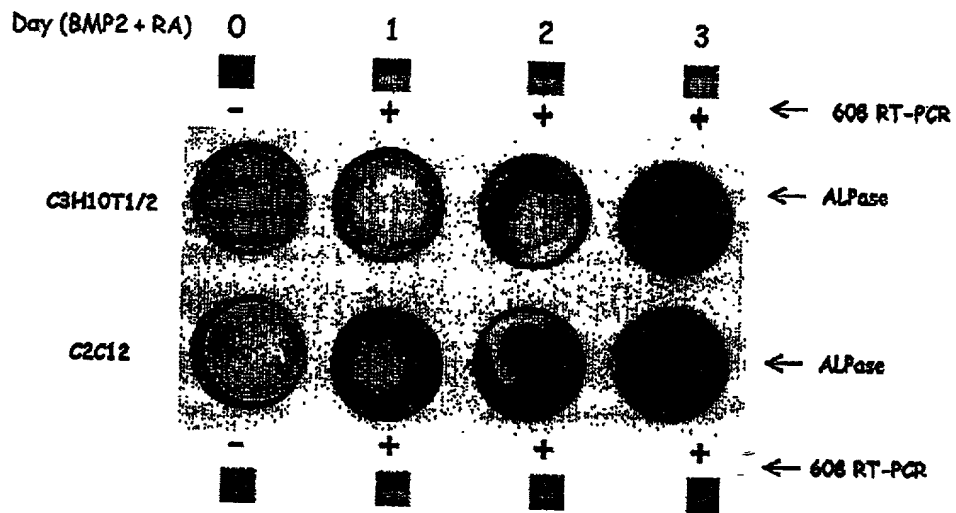


Figure 23

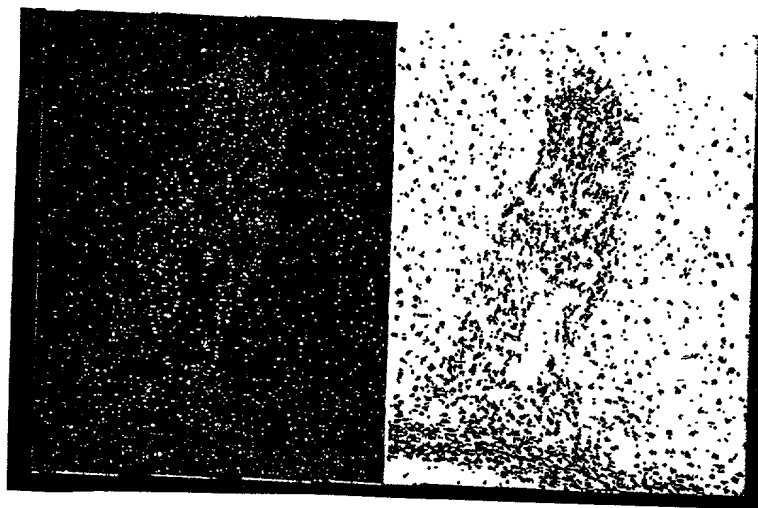


Figure 24

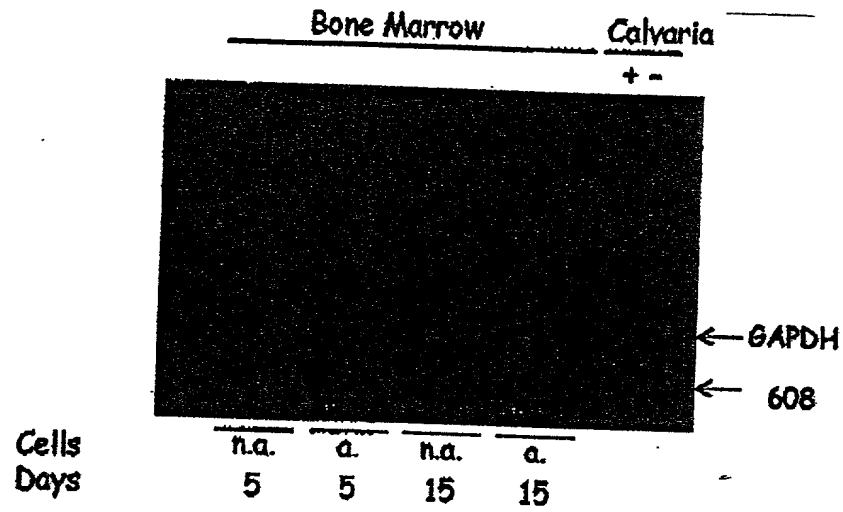
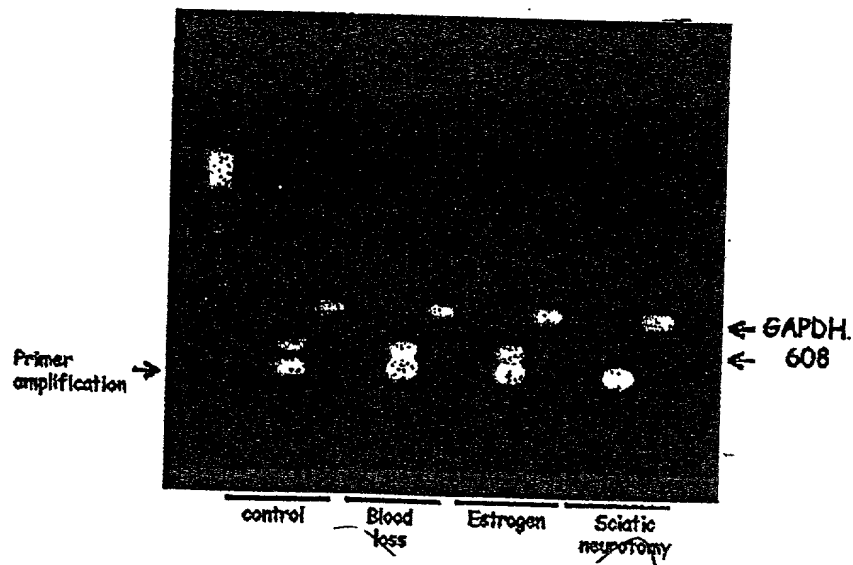


Figure 25



[illegible]

Figure 27

A

B

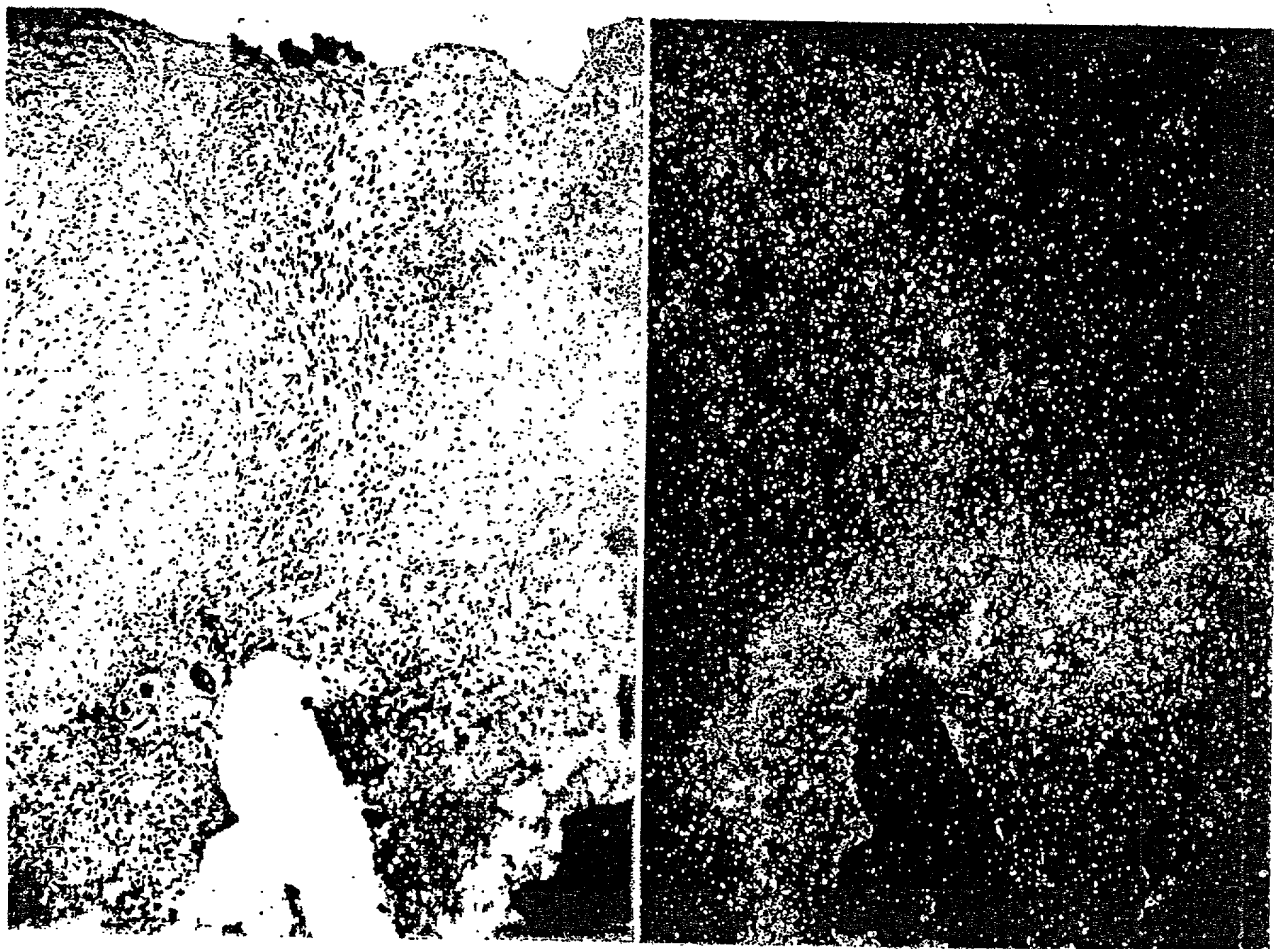
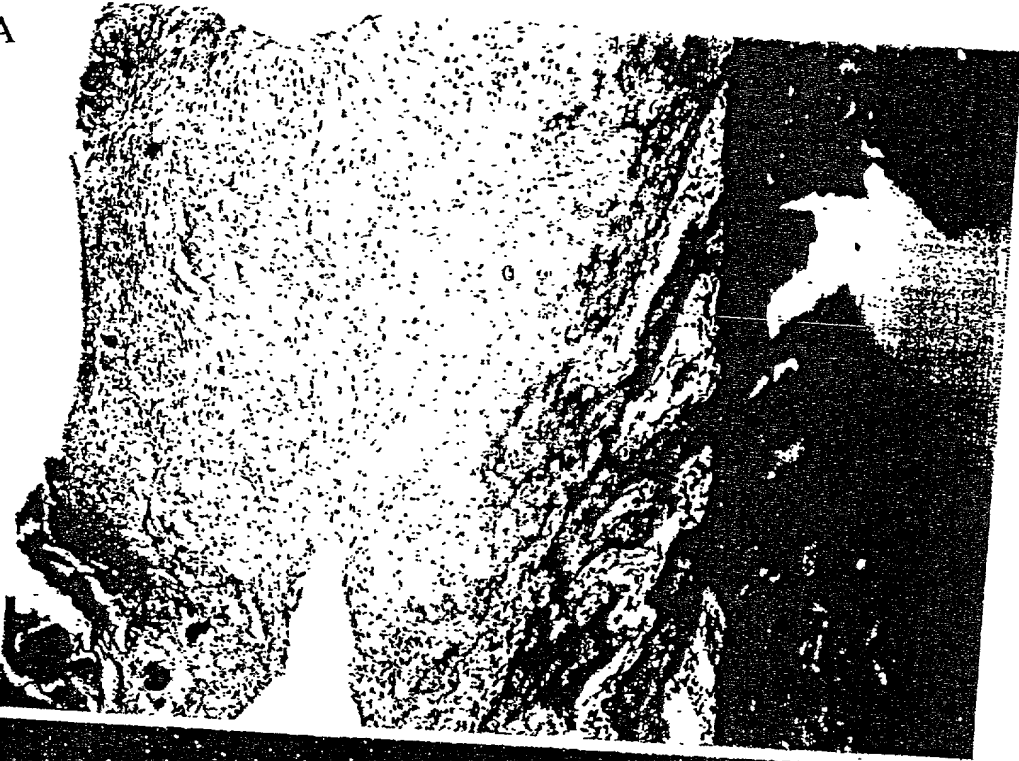




Figure 28

A



B

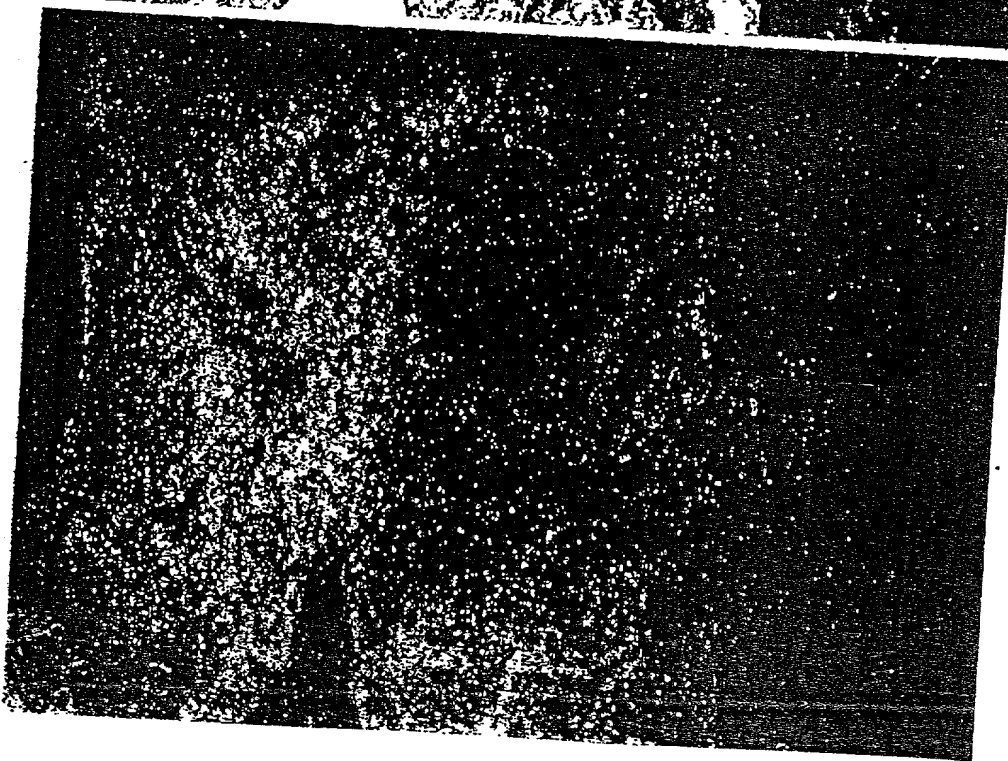


Figure 29

A



B



Figure 30



Figure 31



Figure 32

A



B

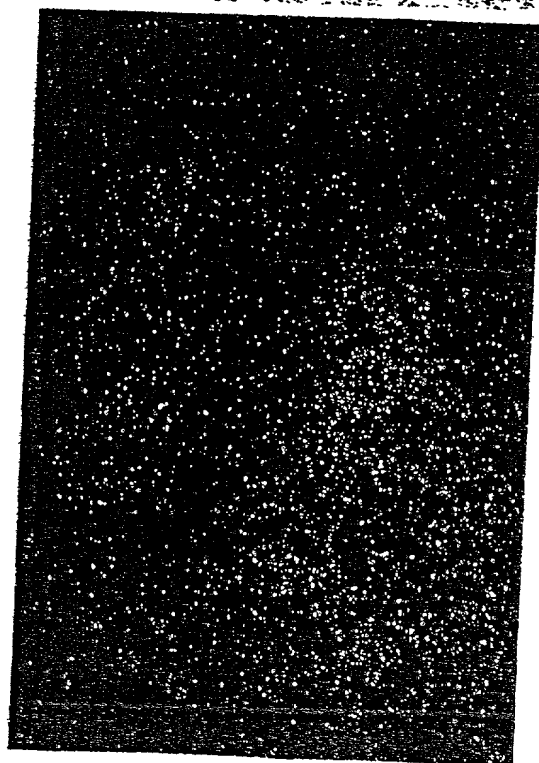


Figure 33

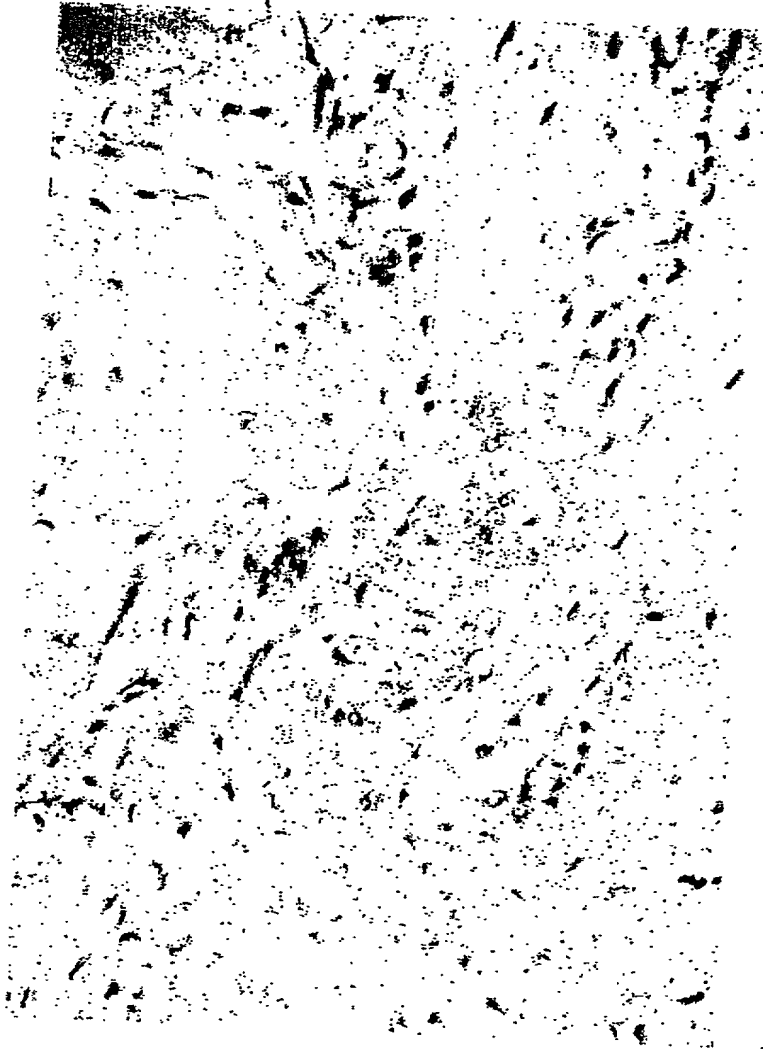


Figure 33

Figure 34

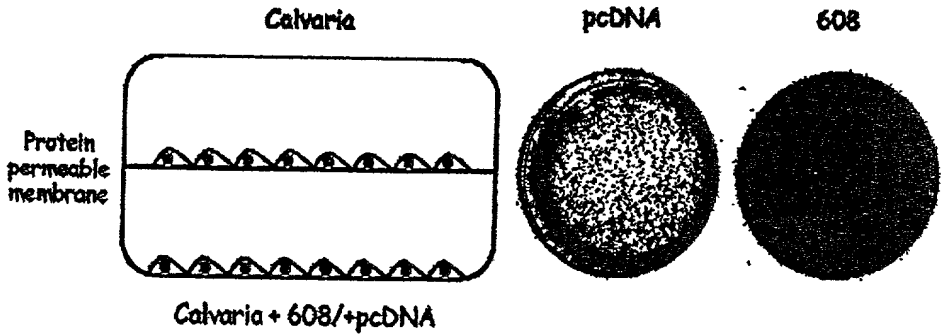


Figure 35

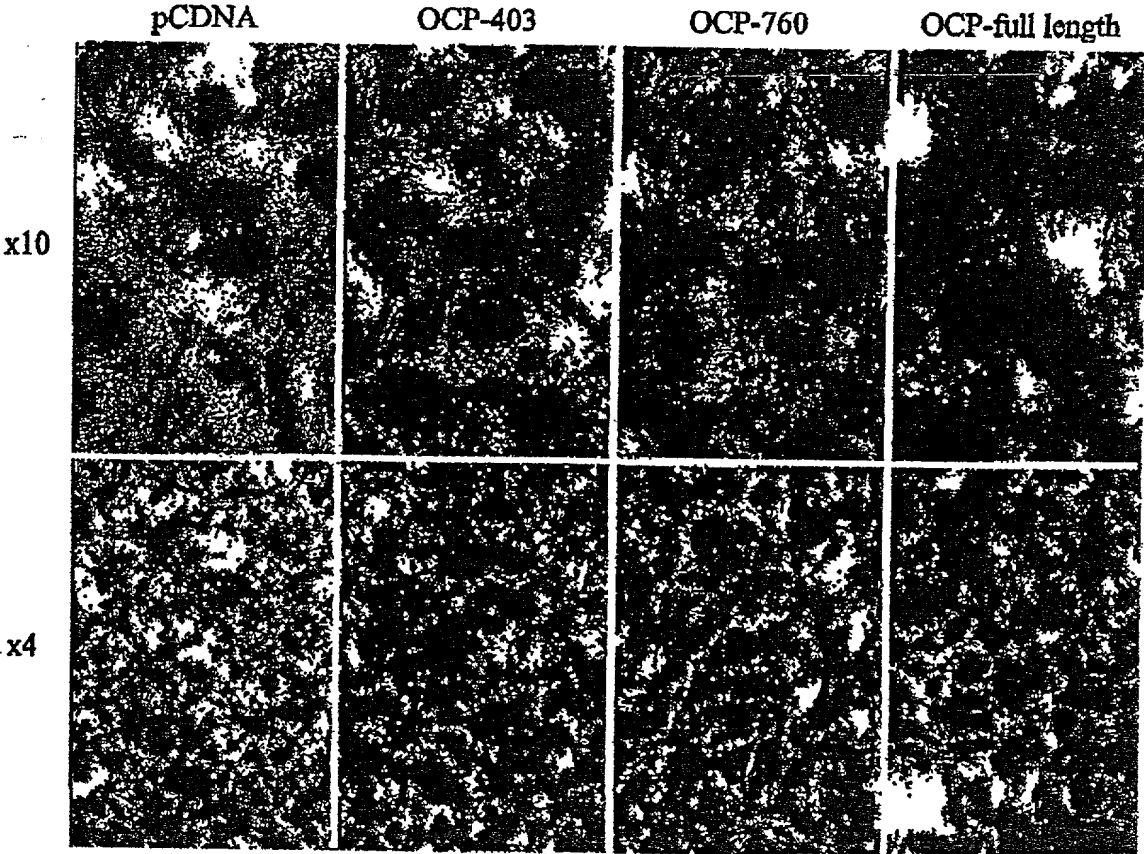


Figure 36

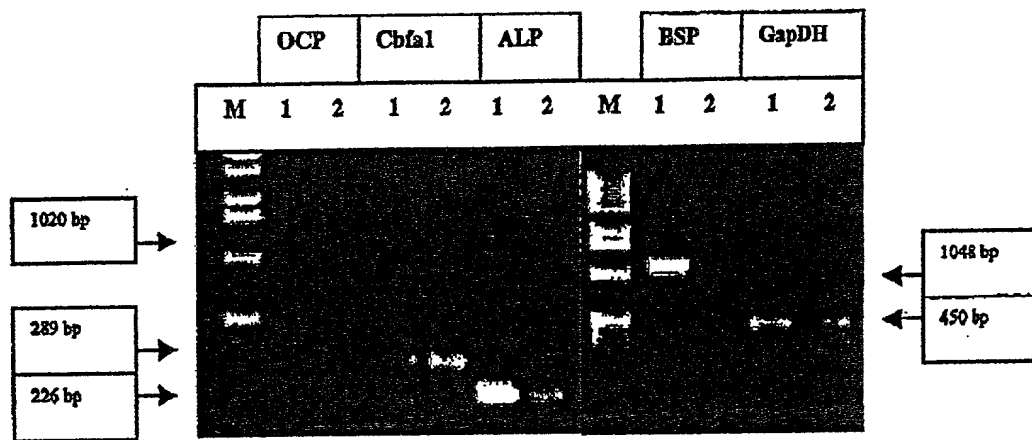


Figure 37

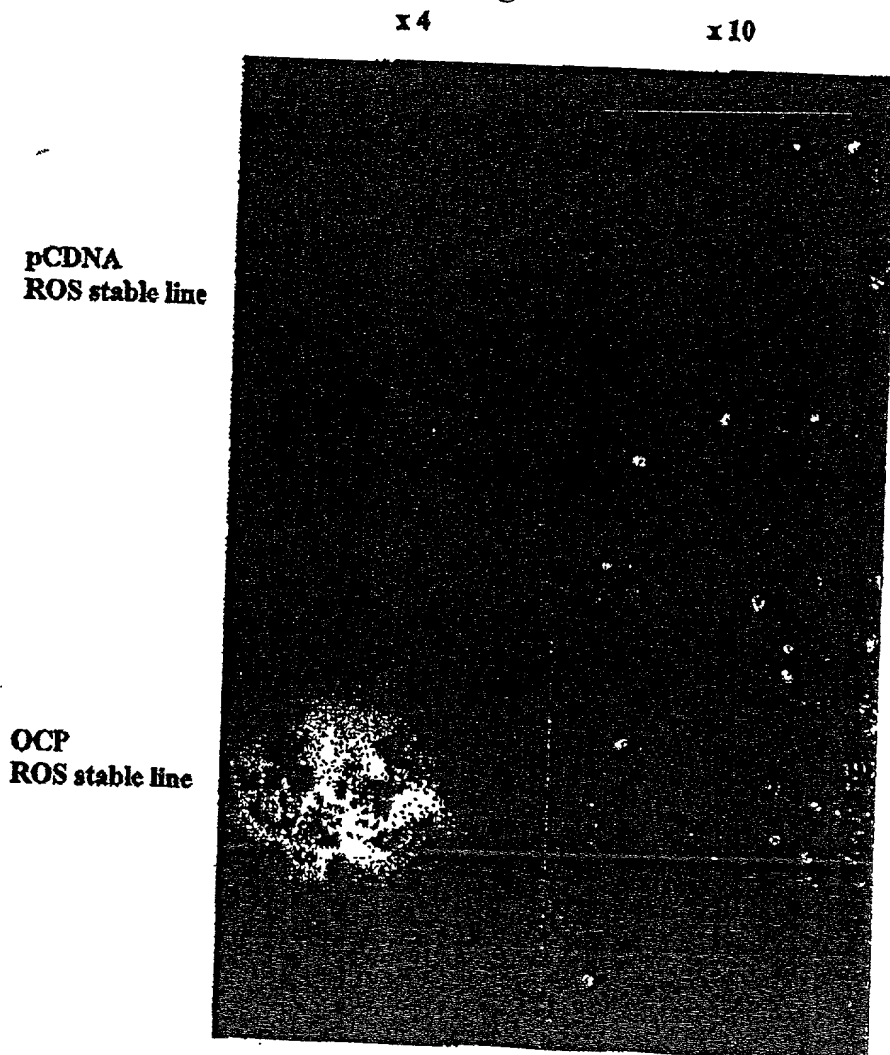




Figure 38

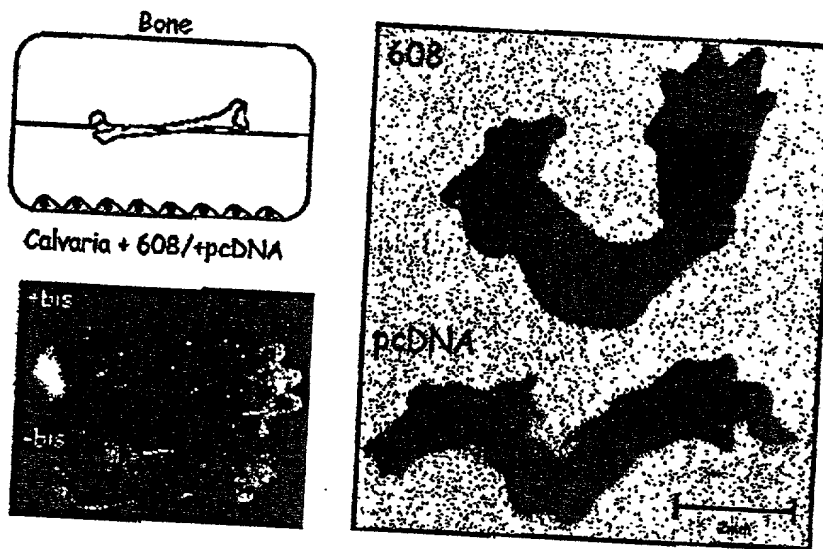


Figure 39

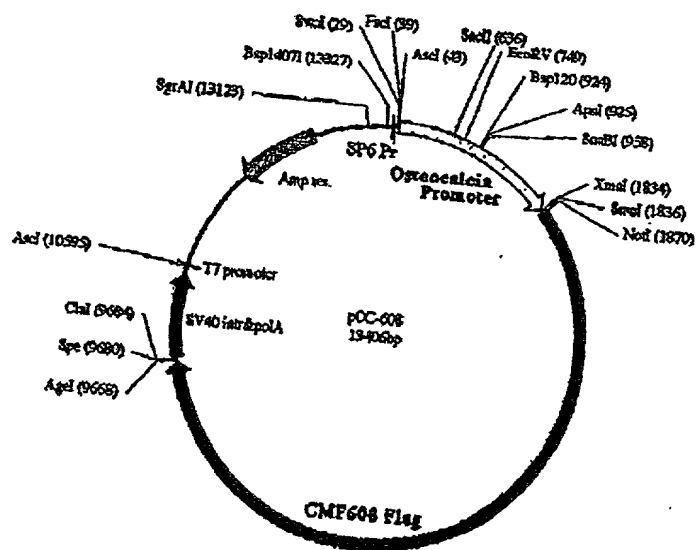


Figure 40

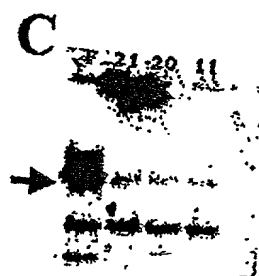
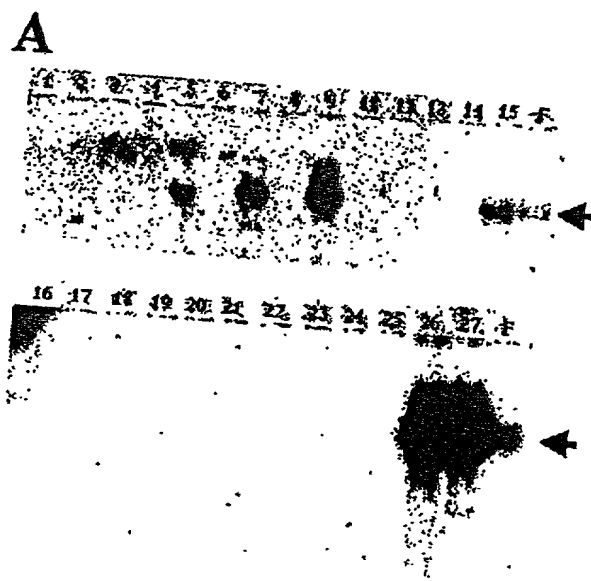


Figure 41

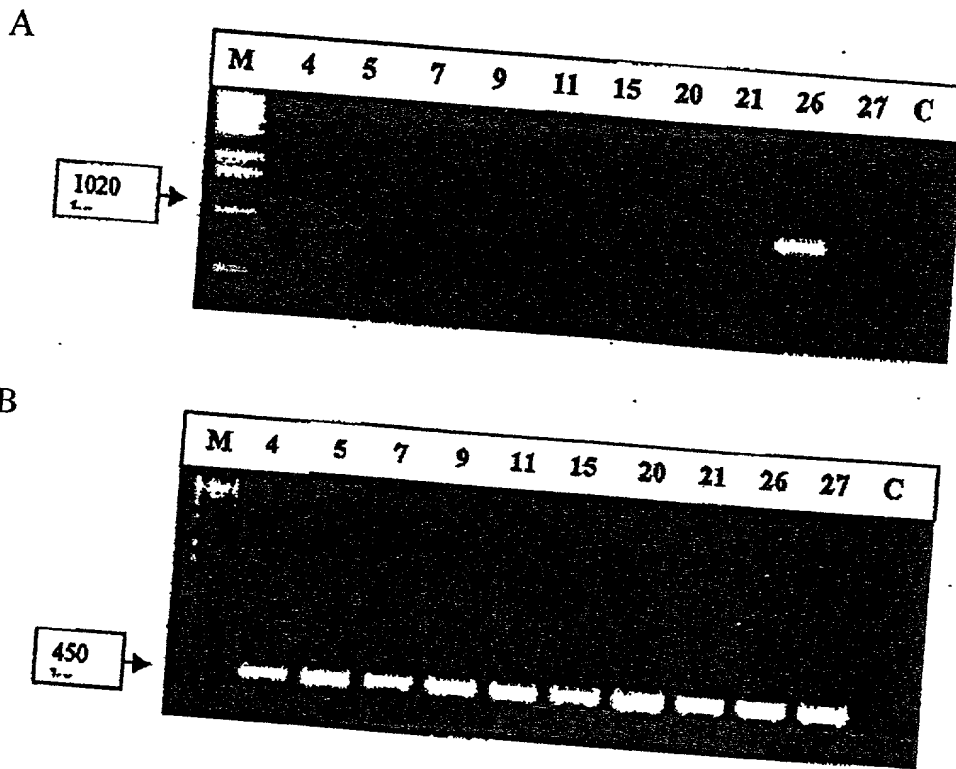
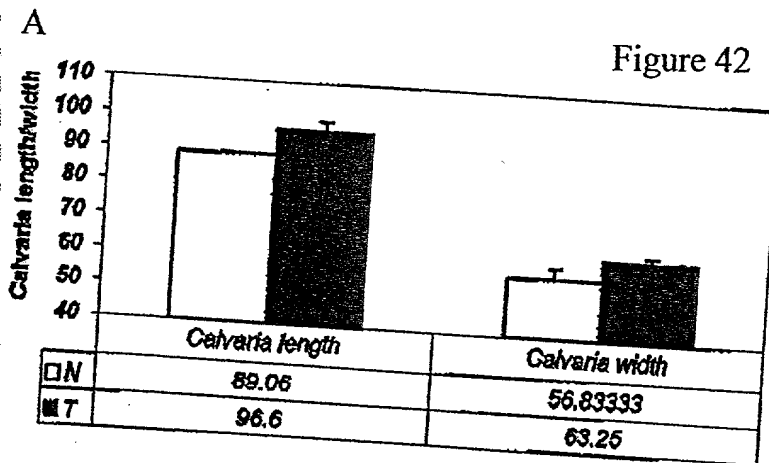


Figure 42



B

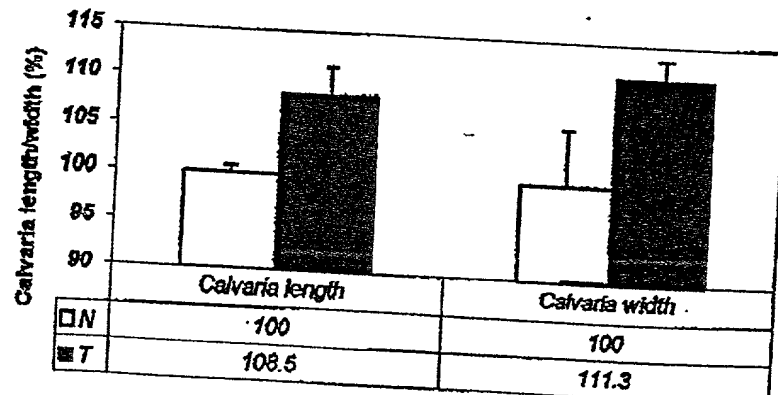


Figure 43

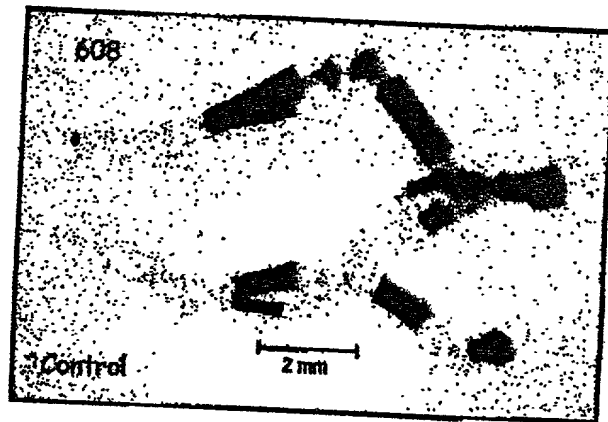


Figure 44

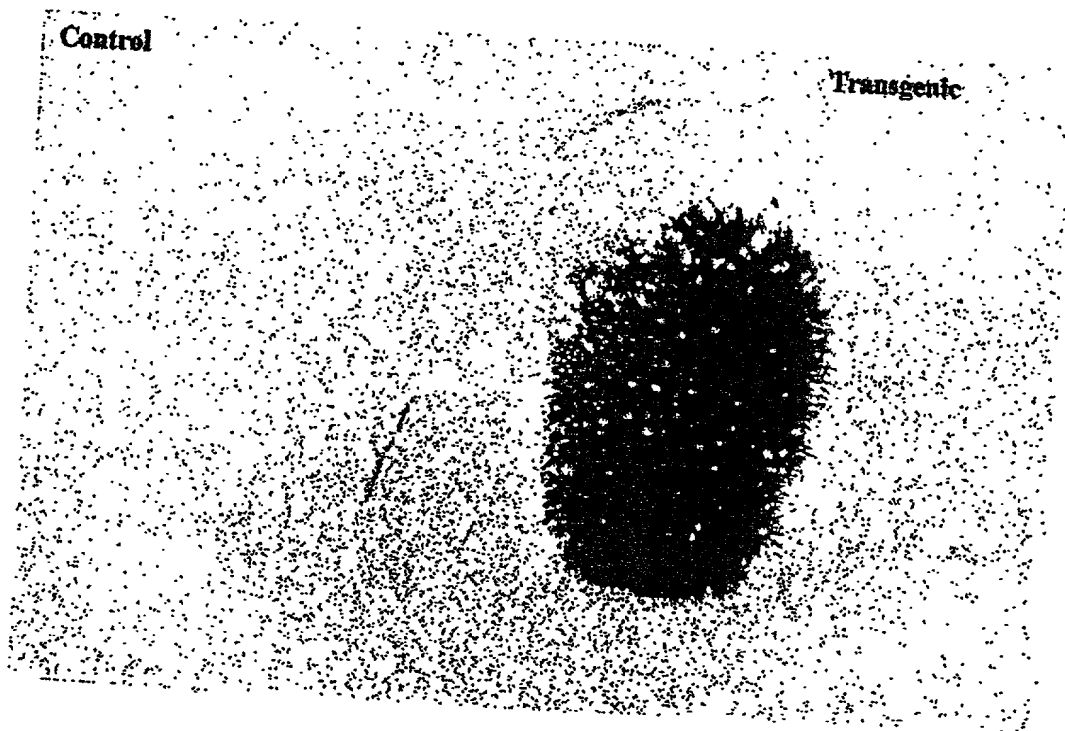
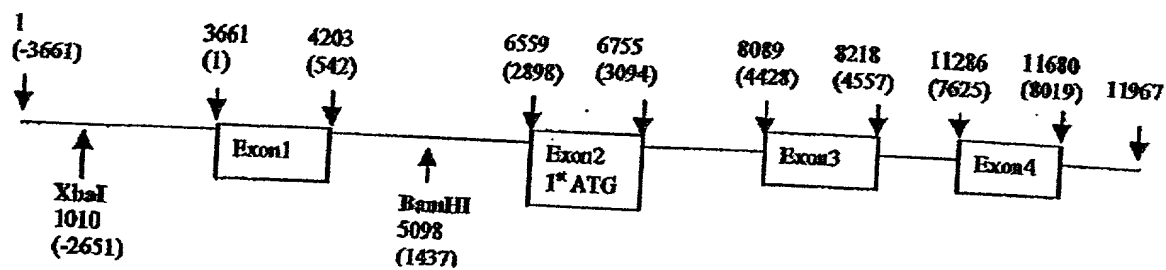


Figure 45



Clone 14C10 /15E11

T7 orientation in pKS

Figure 46

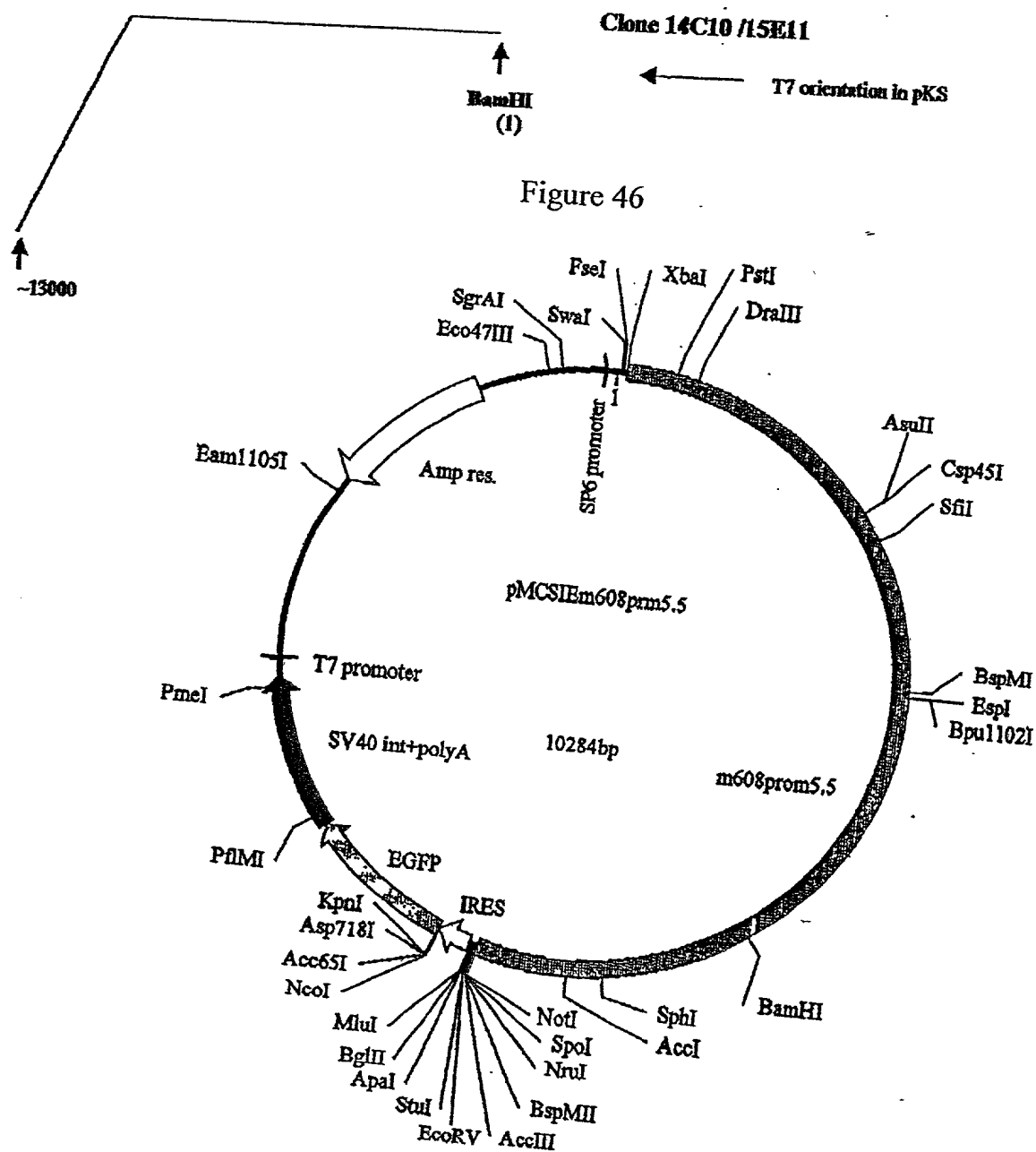


Figure 47

↓ (XbaI)

TCTAGAAAGTAAAAATGATCCTGAGTAGCGATCCTGGGAAAATACGTAAGTCTAACAACACTGC  
AATCATCTCTCTGTGGTTTGGCTGGAGCTGAGGTCTGGAAGGCTCGACCTTGGTTAGAAATA  
ACCTACCGAATACAGAGCTATGACGTTAGTCTGGAAGGAGCTTTGGAAGAATGACAAGCT  
GTAGCTGCCCAGAACATACTAGATGCCATATTTCCAAGGCAAGTGTCCACATGCCGACAT  
CTTAAGAATATGGTTGTCTCTGACGTGCTAAGGACCTTGTTCGTGCCACACAGGTCTCCAG  
GGTTAGTGCTAACTCTGACTGCTTGAATCTTAATTCTACCTTGATCATTAAAGACTAGAA  
ATCAGTTGGTGATTAGCAACTGGATATGGAATATTACTAATTTGTACCCAAGCCAGGCCAC  
CTCAGCTTTGGCAGCTCCATTCTGTGGAGCCAGTCACGTGGGTTTGAATCAACTGT  
ACTGTTTCTACTTACAAGACGATTACCTGAGATGAGTCATTTTCTTCCACAAGTCTTTTA  
GAAGAGTCAATTAGACATATTCTGATGAAGTAAGCATATAAAGTGAGAGCAGCATGAATG  
TGTTCCATGTATGCTCATGGATGCTATTATAATGTGGAATAAACTGACTTTAAAAA  
AGCTTATGATACTTGTACAGAGTAAATCTTCCATAAATATCATCTGCATTATATAAATAT  
TTTCATAATCCATCAATTAAAAACCTTTAGAAATTTTGTAAACACAAAGATCCCTAGGCC  
CTGCCCTAGGATGGTCTGTATGGTGGGCTGAGAGATGGAGCTTAAGAACTTACTTGCTCC  
AGGAGCACATCTTCAGAACATCTGCCCTCAAAACATTTATCCCAATGCTCATCAAAAGGCTC  
ACTCATGTGCTTCAACCACAGGGATTAAACAGTCATTTTAGTCACATTTCTCAAAAGGCT  
GGAAGCCTGTAGAGGAACAGGATGTATCAGGATAACATCCAACCTTACAAAAGGATGTC  
ATAACCTCACCACAACAAACAACGACAACAAACCCATAAAATTTATCAGGCAAAAT  
GAAGTAAAGCATATGCAGAAAAAGTATTATATGTTCTCATTGTGGGGTGTTTTCTTAAAT  
AGTCAATATGCAGAAATATAGACAAAGATGGTTTATGCAAGTGGGGATGGCGAAGGATA  
CTGTAGATTAGAGGACACAAAGCAACAACTACAGAGTGAAGTAATCCAGAGACTTAATG  
TATAATATGAGGACTGTATTAAATAATTCTATTAAAGATACACAGCAAAACGAGTGTATCTT  
ACTAACACACACACTTACATAGAGAGAATAAAGTGATAGATACGTTTGTATTATCTTCATG  
TAGCTGATAATTTATATTGTACACCTCAAAACATAGATAACCAACAAAGAGGAAGAGGAT  
AGGTGCTCTCCAGGGCGGAAGAGTACATTGAAAGTCAGACACCATTTGTGTAGATGTA  
CCACATGGAGGAGCTAGAGAAAGTAGCCAAAGGAGCTAAAGGGATCTGCAACCCCTATAGG  
TGGAACAACATTATGAGCTAACCAAGTACCCCGAGCTCTTGAAGTCTAGCTGCATATATATC  
AAAAGATGGCCTAATCGGCCATCACTGGAAAGAGAGGCCCTTGGAGCTTGCAAACTTTAT  
ATGCCCAAGTACAGGGGAATACCAGGGCAAAAAAGGGGGAGTGGGTGGGCAGGGGAGTG  
GGGGTGGGTGGATATGGGGGACTTTTGGTATAGCATTGGAAATGTAAATGAGTTAAATAC  
CTAATAAAAAATGGAAAAAAG  
TTCACTGCTATCTCACTTGCAAACAGAGGGGAGTCACAAACCCAGGACAAACCACAGT  
GATTGAAGCGTCTTTGAATGTTTATGCTGTTTGTGTTACCAACATCATTAGCATATATTCA  
TGTGAAAACTTACGGGGTCTATGACATGTTTTTTTATTCAGATATATCAGATGCTGTCAGC  
ATATTTGGCACCCTACCAGCCCCAGCCCCCTTTGCCCGCCCCCAACACACACACACACAC  
CACCTTTACCTTCTCCTGGG  
CATCATCTGCTCACTCAACCAAGCTTAATCCTTTTCTTCCCTGCAATAGTACCTCTC  
CTATTTTATGTCTAGGTTCCCTCTCCCTGTTAGGAGATGGGAGAGGTCACGAAAGGAA  
AGAAATTTGTAGCCCTGAGCCAGCCCGGCCACAGAGCCTGCCACACAGACAGGAAAGGC  
CAGGGCTTACCAGCACAGGAGGAGCAAACTCGCAGGCGAGCCTGGGTGGCGCTGGTGGT  
CCCGGGTGGATGGCCCGCCCTTCCAGAGCCGAGGCTATAGCTGCGTCACCTGCCCCG  
CCTCTCTCCGAGTGAAGACCCCTAGAGGCTGAGCAGACCCCAAGGCGGTGCAATTCCA  
TTGGCCCAAGGCAGAGGTGAGCGGCTGCTAATCCCTCGGGAAGTGAAGGGACCCAGAG

Exon1

AGTCTGGTAGATGTGGGAGCTGGGGTTTCAAGGCGAGACAGAGGTTGGGATGGGCAGAAAG  
GGTCCAGGAAAAGGAAAGTACTGGAGGGGAGTTGGGACAAAAGCAGCGACCAAGGGAA  
CATCGCTTCAGTGAAGCCAGGCAAAAGGAGCGGGAAGGATTATATGTAGCTGGGA  
CGCTTTCATAAACACGTGATGACGTGTTTGTGCAAGCAAGCAATTTGAGGAGAAACGGCT  
GGGAGCTCGGAAAGAGGAGTGATCGATTAGTACTTGTAAAGTTTAGGTGAGTTTGAGAAC  
TAACTAACTATACTATTGAGGGAGAAAGGAGCATTCCAGCAGCAGCAGCAGCAGCAGCAGC  
GCAATCAGATAAAGGAAAGCTTTGGTTAGTTTGGAAATGTATGATACCATTAAAAATAACA  
GAAGCGCTCCAGTTCTCTGAAGAGTCACTCCCCAGCTAGTGAAGACTAAGCTACTAA  
GCCTTTTGTCCCGTTGGAAGCAAGAACGTTCTTCAATCAGGTGAAGGCTCTCTCAGA

AGATTTCTCTCTCTCTATGTTACAAGAGGATTCAAAAGCAAGACAGAAGAGCTCAGG  
 TATTGCCAACTCTTTTGTAAATACAGTTTGAGGCTTAAGTGTACGGGAACATCATGTGGTA  
 TTCATTACGGCTCTCTCTCTTATAACTAAGTCTTAAGGTGCATATAGTCTCTCTCTGTTTCC  
 CAGCTACCTTGTACCATCTTGTGTTATCTAATAATAGCAAGCTCATCTGCTTTTAAATCATC  
 AOCAGAGAGTATTCAAAAATATTAGTGATGTAACAGTGACAGTGTAGGCATAGAAAGTA  
 ATCATTAGTAAATCTTAATTTGGGTTAAACTCATTTCATAACAGCTCCAGGTTGGGAGGGAT  
 CACTGAGCCTTCGCCACGTGCGGGTTAAAGATATTTTCTAACAAGAGAAGCAGAAATCTTTC  
 CTTGGCCATGCTCCCCATCACTGTGTAGTAAGCAGAGGGGTGTTTCCAAGCAGAGAAAAG  
 AGCAGACAGTGTATGCTGCAAAAGTCAGAGACTCAGCCCTCCAGCTGGTCAGTTTACT  
 GTCTCCCGGTCATTAGTTGGCTCTGAAAAGGCCCATGTGTCTTATTGGCAAGGACTTGC  
 AGACATGCTAGAAAAGAAATTTGACCTTTTTTCTAGTGGGTATTACAGCTGTAAAAGTAT  
 TTTGGAAGGTAAAGCCAAATAAATAAAACACATATTAAATAATACAATGTTACAAAATTT  
 GATCATATAAAGAAAGTACATTTCATAAATGCAATGTGAAAAATATATATAATTTTATCTAT  
 TTTACTGGTGCAAAAGTTTTCTAAATTTGCACATGTACTATTTTATATTATAAAAAATTTTT  
 AAAATGTATATAAAAGTGTAAAAGGCTCTGGTCAAACAAGAGAGTTAAATTTACAAACT

↓ (BamHI)

TTAATTGTCCCGATAACATTATTATGATCTCTAATGACAGGGATCCTGCTTTTCATTGGGA  
 AATGAGAAGCTATGAAGATATGTTTACAATAATAAGCCATTTAGTGATAAAGTCCAATG  
 GGAAGCTAGCACACACTGGTTTATAAAGAGAACAGTTTCTGAGTCTATGCAAGTTTACA  
 CTCTAGGGAATAAGAGTTCTCTTTCTCCAGATTTCACAGCATTGTTGTGTCATCTTATC  
 TTCTTGATGATGAGCATTATAAGTGAATAAGATAGGATCTCAAAGGAATGTCAATTTGG  
 ATGCCCTGAACAATCTTTTCAGGTCTTTCTTTCAGTTCACTAGTCTATTCTTTATGGATAA  
 TTGGGGGATGGTGTTAATTTTTCAGTTCTTATGGAATTCAAAAAACAAAAACAAAC  
 AAACAAACAAAAACCTCTGAACTAGAACTACCAATCCATTACTGGGTATGTAACAAAG  
 AGAAATCTGCACAGAAATTTATTGCTACATTGTTTATTTCACGACAGCCAAGAAATGTGGA  
 ACCAATCTACGTAGCGCTCAAAATATGAACGGATAAAGAAAAATGTGGAAATGTGTACAAC  
 AGAGTCCCATGTGGCCATAAAGAGTGAAATCATGACATATGCAGGAAATGGATGCAACT  
 GGAAATCAATTTGGGCTAATCAAAACAAGACAGACTCAAAAGGAAACACCGTGTAGCTT  
 CTCTGACAAACAGAAGCTAGATTACACTTGTACGTGCGCATGTGTGTTAGAAATTTATT  
 TAGTTATACACTATTCTAATCTGTGAGTGTGTATAAAGGCATGCATGTAAAGCAAAAAACA  
 AGCTAGCTGGGGTGGGTAGGAGAGAAAGCAATGAGAGGAGTTAATAAGAACGAAGCATA  
 GTAACATAGGTGCCAGGATGAAATGCATTAAATTTGTATGCTAACTAAACACAGACAGGA  
 GGCACACGTTCAAAACAGGGTGAAATCCACAGCAGAGAAAGGGGAAGTAGACACAAAGT  
 TTCGCCACTAACCAAGAAGCCATTTGCAGTTGCTGCTGCTGGGAGGGGCGTTCCAGTTTT  
 CTCAGTCTGACACTGTGTATAACAACAGTTGACAATACAAAGTTGGCATGATGGATGG  
 TTTTGTGCTATTTTTCAATTTTTTTTCTTACTGTTTTGTTGTTGTGGTGGTTGTTGTGGTGGT  
 GCTGTGGTTTTTCAATTTGTTTCTTTTGTAGAGAGAGAAGGAACATGAAATTTGGGTGGGTAGGA  
 AGCTGGAAACGATCTGGAAGAAGTTGGGGAAGAGAAAAATTTGTATGGAGCATATTTAA  
 ACAACAAACAAACAAACAAAGGTTCAATTTGCCACAAAAAGGTGTGAATTAATTAAC  
 TAGTTACGACTCTTAAAGAAAAATTTCCCAATTATTTCCAGAGTTGCTATGTATGCTGTGC  
 TAGGACTTTGCTTGAAGTGGCCCTATAACTCTGGTGTGGTGTCTTTTCAGG - Artificial

Not a site

(SEQ ID NO:17)

Figure 48

CACAGACCTTCCTCTTCTAACCTCTCTCCCCATCTTGTGCTTCATCCCAGACTTCAACAC  
CAGCAAGCACACTCTGCTAATGCAAGGGCTGCTCCTGTGAGGACAACAAGGAGGCTGAAG  
GCAGACCCACACGTTTCCAACCTGCTCCTGAGAGTCAATCCCCCTAGACTCATCTATAGCAG  
GAAACCTGCTGTGATCTCCATTTCTTCTGACCAATCCCCAAGTTATCACAAGGAGTTT  
TTCCTCAAACCTTTCTCTCCAGCAAAACCCCTTCAGCTCCTTGGGTACTTTCTCTAGCCCT  
TCATTGGGAACCCCTGTGCTCCATCCATGGATGGCTGTGAGCATCCACTTCTGTATAGAAT  
CTTGGTCAGTGCAGTCTTTTGTATCCTCAAGAACACTGGGTCTGAAAAATTTAAACCCAAAG  
AACTGTTTTTGTATGATTGCTGCAATCTCTTCAATTCCAATAAAGAGTAAGCATCTCAT  
TCCTTTGTCTCCTCCTTTTCAGTACCACCCCTGCTTTGTCTGCTTTCTCAAAGAATCAATAAA  
ACCAAAGTGATATAGATTTCATGGCATTCTCTAACTGCTACATCCACTCCAGTAGTATCTC  
ACTTGGCAGGTGTAAAAGCCTGGAAAGCAGTCACGAGGCAGTTTTCACAGAAACTTAGCCCTC  
CTGGAACCTTGGCATTCCCATAGCTAGAAATGCCCCAGATTGTGCCCTGAGATATTGTGGTG  
GGTCTTGCATGCTTTCTTGCAGTATTTTACTGGATAAGAGTTAGAAAATCTCAGGGCGAGCT  
TAGCAAAAGTATACCTAGAATCTTTCATGACAGTCAGGTATTGCAAACTACATTGCATATTA  
GAAGAAAAGTTGGTAAATTTCTTCTGACAAATGGAGATTCCCTACAGATAACTTAAAGAAC  
AGCTAAGTCACACTCATATGCAAGAATTTACCAAGGCCTAGGAAAAGGGGGGGGGTACT  
GCTTTATTCATGATAAGGTCTGCTAGAGCAGAACCCCTGGTGTAGCTTTCACAAGGTTTC  
AAAGGTGTAGCATAAATTTGTGACTAGAGTGTGAAATCTTTACCTGTCTATTAGCTGACTCTA  
GGCAGAGCTGTTTTATCTTTACTGTAAACATTACCTGGTTCTGTCTCAGTCTTTGAAGGCAT  
TCCTCTGTTTTGTGACAGATACTTCTATGTACCTCGCTGTGTGACACCCCTACTCCTTTGT  
TTTCTGTATTATATAAGCCTGGTGTTCCTTTGTGAAAAATTACATCCAGATACAGCACTC  
CCTTGTGTCTGTGCTCTTTGTCTATTCTTGGCCAACTCCATGCCAACCTGCCAGAACCCCTA  
GTCTTTTCCACAGATTGAGGGAGGCCGACTGAGCCTGGTCCATGGCATCTAACCACTGTCA  
GCTCACTGTTGGTGACTACCTCAAGGTACAAGCTCCATTACTAATGAAACAAAATTAGAT  
AAGTGTGGGTCCAGGAAGCAGGTTGTACACCCCTGTCTGAATGAACATTATGAAATGACTG  
AAATAAGTTAACCCATCTCTTCTCGTTTGTCTAATATAGCAATAAACCGAGTTTCTGAGC  
TGCTGTGCTGTGTCTCTCCATCAGAGGGCAGAGCCAGTCTGATCTCTAGCTTTCTGTATGTG  
TGTCCATTGTTTCTTCAGTTCCTGTGTGCCCAATTAGGAAATCCTAAGCCATGAAAGCCATG  
AATCTGGGAATGACTTTTCTAAGAAATGCCACGTGAACCTTGCCTTTCAACGTTTTGCCTG  
TAAACAAGATATATGGTGCGCAGTTTATAATCATAATAAGCTTTGAAATAATATATAACTC  
CATTCTCATTCTGCTTCCACGCTGAGCATCCTGTTTCCCAGGGACCACAAGAGCATTGTA  
AAAGTAGTGATTATGACCTGCTTTGTCTGTACTATAAAAGCTTTCATGAAAGGGCAGCC  
ATGTTGAAACATGGAACCTGGGGTGACCTGTATCTGTGTTTCTGCGTGTGCTCACTCATA  
TTTGTCTCCAGAATAAATGAGTTTATCAACTTCGAGGAAAAAGTTGTGTGTTGTATAGCA  
CGCCCGTGGAGTCCCAACCTTCTACTTCTGTAAATCTGTATATGGTAGAAAAAGTTAATTT  
ATGTGATTCTTCCAACCTCCAAATATTTCAAACTTTTAGCCCTCAGCCTGGGATTTCTTTG  
ACTAAGTCTATTGATTGGAAGATCTCAGTGGTTAGGATTTGCAGTCATGATGTTTCATACG  
TCAGGCTAAGCTGAAAAATATGACAAATGAAATGTCAAATGTCAATGTGCTGCTGGGAATGTG  
AGTGTAGGGGGTTTTAAAGAAACAAATACCTACTCTAAATAGTTAATAAGTCCCATGGTT  
CTATTCTAGTTTTGAATAATGTTCCCTAGTATACAGCAATTTAATTTGAAATGAATAGCTTC  
TTATCTTGACCAATCTCAGTGACTTCATCCGTCCCAAGTCATGTTTTTCATATTCTAAGGAT  
AGGTCTCATTCAACCACATGTTTATCATTGGGATCTGCATTTTTCTGTATGCAAAATGATTT  
ATTCTTCCAGAGCACTGGAATTGGGTGAAATCATCTTATAACGGCCAAAACCTAAATGCTTC  
TGTGCTAAACAGAGAGTTACAAGACCTTTTTATGTGGATGGCAGCATTTTAGTCATCCTTA  
TGACAGAATGTGAGAGTGGAGCTCCCACTGGGGGAGGGGCTGGTCTTGGCAGGATTCTC  
TTGGAACATCACACAAAGAAATTCCAAATATGAAATGCACATGATCCATCCAGAATGTG  
ACTTTTGACTCTTGAACATGAGCTTTTAAAGTACGTTTGGCTGTTTCAACCTTGACTTTGAG  
GTGAAGGAAAGCTCGCCAACTCCTTTTATATGTAACACAATATATCAAGATCTAATGTGA  
GACAGTATGCCAGTCCCAAGATCTGTCAATATGACTGAAGACACATTGGGATGTTATCACT  
AAGGCAGGAGAAGGCAAGCTACAGTGAAGCCAGTTCATATAAAGCTTTATGAGAAATT  
AGATAAGAAGGGTTTCTAATTTTAAATTTTTTTTATTAGATATTTCTTTCATTTCATTTTC  
AAATGCTATCCCAAAAGTCCCTTACTCCCCCTCACCTGCTCCCTACCCACCCACTC  
CCATTTCTGGCCCTGGTTACTTGTGATAGTGGTCATATGATCCACCAAGCTTTACATGCTC  
ACTATCTGGTCTATTGCAAGAAATGGCTGCCGAGCTGATGCAGTCAGATACAGACACCTAC  
AGCCAAACAGTGGAAGGAACTTGGGGACTCTTATGGAAGAAAAGGAGGAAGGGTTATGG



CCCCGGATGGGGAAAGGAACTCCACAGGAAGACCAACATACTTGGTCAACTAACCTGG  
ACCCTTGGGGCTCTCAGAGTCTGAACCACCAACCATAGAACATTTCATGGGCTGTACCCAG  
GCCTCTCCACTCATATGTAAACAGATATGTGGCTTGGCCTTCATCTGGGTCCTGAACAATA  
GATGGGGGTTAGGGGTGGGGATGGGGGTTATCTCAAAAGCTGTTGCCTGTATGTGGGATA  
TGTTCTTACTGAGCTGCCTAGTCTGGCCTCAGTGGGAGAGGAAGCACCTAGCTTTAGCCTT  
GTAAAGACTTGAAGTTCTGAGGTGTCCGTGGAGGGTATACTCAGGGAGGCCCTCAOCTGC  
TAGGAAGAGAAGAGGAGGGGGAAGACTTGGGGGAGGGGGCAGTGAGCAGGTTGGTAAA  
ATGAATAAGAAAAAAAAAAATAAAAAATAATTAAAAAAAAAAAAAAGAATGGCTGCTGAGCCC  
TACTCTAAAACCATTGCATCCCCCCCCCAATCATTCACTGACTACGAATTAATCATT  
GATACTAACAATAGATGTAGGAACTATTGTTAACTTCTTTGTGAECACGAGTGGTATTTG  
GAACCTTTTTTATTGAAGCTTTCACACAGAGCCTTGTCTTTCATTTCCCTGTACATGCATG  
TAGCTTAATGATGTTTCAGTGAATTAATAAACAATGAAGAATAAAGACAACTGTATTTT  
AAGGATTCTTCGTATATATTTAAAAATCTAAGGTGGTCAOCTGGAAGAAATGTCTTCAGTT  
TTTCTATATATGTTTACTCTATCGTATGTTAATTAATTATATGCAATAATTCATAAAATCTA  
CAACATAGTATGTAACCTTATAAGAAAAGTAAAAACATTCATGAAATTGTGAAGGTTACTTTTC  
CTTACCTCAGAAACACTGGGTTTGAATAATTTCTATTTTGGTATCAGTGAAGAATTTGAA  
AGAATGTAATAACCTACTAAGGCAACATAGAAGTTGAAATTAAGAGAGTAGACAGGA  
GAAGTAATAAGGCAAAATAATGAATATTGCTTTAAATAGTTCTTAATGTATCATCTAACTA  
GGGTGTGATTCTCCAGACTTGACTCCATCCAAAATATCCAAAATGACTCTAACCACAGTCA  
TTGAAACAATGTGTTGAAAATAATAAACATTTCTACTTGAATTCAGATTCTCCTACT  
TTGCTTTTATTGCTGTGATAAGCACCATGAACAAAGCAGCTTAT

(SEQ ID NO:18)

Figure 49

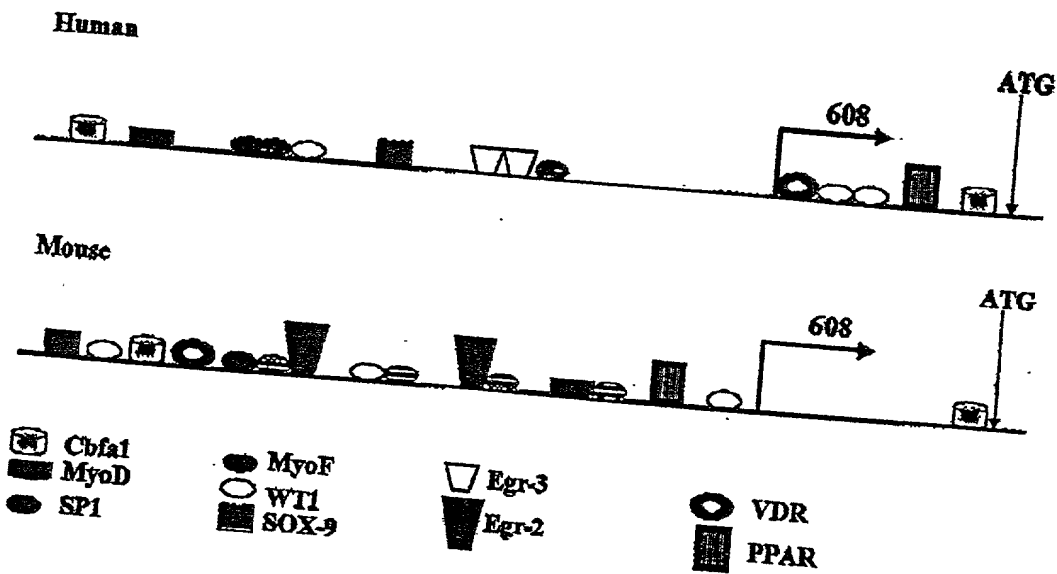


Figure 50

TAAGCCTTTTGCTCCCGTTGGAAGCAAAGAACGTTCCCTTCAATCAGGTGAAGGCTCT  
CCTCAGAAGATTTTCATGTCTCAGCTTATGTTACAAGAGGATTCAAAAGCAAGACAGA  
AGAGCTCAGGTATAGCCAACTCTTTTGTTAAATACAGTATGAGGCTTAAGTGTACGG  
CAACTCATGTGGTATTCATTTGCGGCTCTCTTCTCTTATAACTAACTCTTAAGGTGCA  
TATAGTCTCTTCTGTTTCCCAGCTACCTTGCACCATCTTTGTTTATCTAATAATAGCA  
AGCTCATCTGCTTTTTAATCATCACGCAGAGAGTATTCAAAAATATTCAGTGATGTA  
ACAGTGACAGTGTAGGCATAGAAGTAATCATTAGTAAATCTTAATATGGGTTAAACT  
CATTCATAACAGCTCCAGGTTGG (SEQ ID NO:19)

TTTGAACCAACCCAGATGCCCCTCAACAGAGAAATGGGCCAGAAAATGTGGTCCA  
TTTATCCAATGGAATACTACTCAACTTATTAACCAACGACTTTCATAAAATTTTAA  
GGCAAATGNATGGTCTGNAGGATCTTGAGTGAGGTAACCCAATCACAAAAGAACAC  
TCATGGTATGCACTCACTGATAAGTGGCTATTTGTCTATGGAGTGATTTAAAAGGGA  
AGAAGACACATAGCTTTTTGTGTGTATAATATTAAGATGGAAATTTGCCAGTGCTGT  
TTGGCTTATGAGTGAATCTTGTTTCAGTGGATTACCGGAAGAAAATAATAAGTGAAC  
TGTAGGAAGAAGTAGTTAATCAAGGTGACAAAGTATCCTGACACATTGGGAAAAGA  
CCACAGTCCAGGAACTGAGTCTTAAGGATTCATATTAAGTCCAGTTCCCCATGTGC  
AGCTCTGAGACTTTGGCAGATCAGACACTTAAGTTCACAGCTTCCTACACAGAGCA  
GTTACTATCCTTGCCCTTACACATGGAGTGTGCCATTAAGTGCCTGAACATGAGTCT  
GACTTGTTAATAATCTTTAAATCCAATTGTGTGTAAAGTATGTGACCAAAGAGCAT  
GGTCATGCTATTAACCTTTGATGTTCTATGGACTCTTAATTTTATGGTAGAAATGTCA  
ACAAGCTTGTGGAGGCTGGAAGATACAAGGCTTAAGAGGATGGCCTTTCAGTTTTG  
AAAGTAATTCAGTATGTGTTCTGGCATCCCTTTTCCTAAAGCAATTTAACCCCCAA  
GTAGGCATAATTTTAATGCTTACTTCATCAGAATATATCTAATTGACTCTTCTAAAAA  
GACTTTGGTATGCATAGGATCTAAATGTAAATGTGATTTACTGACATAATAAATAGG  
AGAACTGAGCTAGAATAGGTATAAAATATGTGCTGGCTTTCTAATAGGTCTTATAG  
GTTATATAAGAGGTGGGAAAGGAATATTTGAAACATCTAGAAGTAAAATGATCCTG  
AGTAGCGATCCTGGGAAAATACGTAAGTCTAACACACTGCAATCATCTCTCTGTGGTT  
TGCTGGAGCTGAGGTCTGGAAGGCTCGACCTTGGTTAGAAATAACCTACCGAATAC  
AGAGCTATGACGTTAGTCTGGAAGGAGCTTTGGAAGAATGACAAGCTGTAGCTGCC  
CAGAACATACTAGATGCCATATTTCCAAGGCAAGTGTCCACATGCGGACATCTTAAG  
AATATGGTTGTCTCTGCAGTGCTAAGGACCTTGTTCTGTCACACAGGTCTCCAGGG  
TTAGTGCTAACTCTGACTGCTTACTCTTTAATTCTCCCTTGATCATTAAATGACTAGA  
AATCACTTGGTGATTAGCAACTGGATATGGAATATTACTTAATTTGTACCCAAGCCA  
GGCCACCTCAGCTTTGGCAGCTCCATTCACTCTGTGGAGCCCAGTCACGTGGGTTTG  
AATCAACTGTACTGTTTCTACTTACAAGACGCATTACCTGAGATGAGTCATTTTTCTT  
CACAAGTCTTTTTAGAAAGAGTCAATTAGACATATTCTGATGAAGTAAGCATATAAAG

TGAGAGCAGCATGAATGTGTTCCATGTATGCTCATGGATGCTATTATAATGTGGAAA  
TAAACTGACTTTAAAAAAGCTTATGATACTTGTCACAGAGTAAATCTTCCATA  
AATATCATCTGCATTTATAAATTATTTTCATAATCCATCAATTAACCTTTAGAAA  
TTTTGTTAACACAAAGATCCCTAGGCCCTGCCCTAGGATGGTCTGTATGGTGGGCC  
TGAGAGATGGAGCTTAAGAACTTACTTGCTCCAGGAGCACATCTTCAGAACATCTGC  
CTCAAAACATTTATCCCAAATGCTCATCAAAGGCTCACTCACATGTGCTTCAACCAC  
AGGGATTAAACAGTCATTTTAGTCACATTTCTCAAACGGTGGAAGCCTGCTAGAGGA  
ACAGGATGTATCAGGATAACATCCAACCTTACAAAAGGATGTCATAACCTCACCA  
CAACAAACAACAACGACAACAAACCCATAAAAAATTATCACGGCAAATGAACTAAGC  
CATATGCAGAAAAAGTATTATATGTTCTCATTGTGGGGTGTTCCTTAATAGTCAA  
ATATGCAGAATATAGACAAAGATGGTTTATGCAAGTGGGGATGGCGAAGGATACTT  
GTAGATTAGAGGACACAAAGCAACAACACTACAGAGTGAAGTAATCCAGAGACTTAAT  
GTATAATATGAGGACTGTATTTAATAATTCTATTTAAGATACACAGCAAACGAGTGT  
ATCTTACTAACACACACTTACATAGAGAGAATAAAGTGATAGATACGTTTGT  
ATCTTCATGTAGCTGATAATTTTCATATTGTACACCTCAAACATAGATAACCAACAAA  
GAGGAAGAGGATAGGTGCCTCTCCCAGGGCGGAAGAGTACATTCGAAAGTCAGACA  
CCATTGTGTAGATGTACCACATGGAGGAGCTAGAGAAAGTAGCCAAGGAGCTAAAG  
GGATCTGCAACCCTATAGGTGGAACAACATTATGAGCTAACCAGTACCCCGGAGCT  
CTTGACTCTAGCTGCATATATATCAAAAGATGGCCTAATCGGCCATCACTGGAAAGA  
GAGGCCATTGGACTTGCAAACCTTTATATGCCCCAGTACAGGGGAATACCAGGGCCA  
AAAAGGGGGAGTGGGTGGGCAGGGGAGTGGGGGTGGGTGGATATGGGGGACTTTT  
GGTATAGCATTGGAAATGTAAATGAGTTAAATACCTAATAAAAAATGGAAAAAAA  
AAAAAAAAAAAAAAAAAAGGAAGGTCAGACACCTCACTTCACTGCTATCTCAACTTG  
CAAACAGAAGGGGAGTCACAAACCCAGGACAAACCACAGTGATTGAAGCGTCTTTG  
AATGTTATTGCTGTTGTTGTTACCACCATCATTAGCATATATTCATTGTGAAAACCTA  
CGGGGTCTATGACATGTTTTTTTATTCAAGTATATCACATGCTGTCAGCATATTTGGC  
ACCACTACCAGCCCCAGCCCCCTTTGCCCGCCCCCAACACACACACACACACAC  
ACACACACACACACACACACACACACACACACACACACCTTTACCTTCTCCTGGGCA  
TCATCTGCTCACTCACCCACCCAAGCTTAATCCTTTTCCTTCCCTGCAATAGTACCTC  
TCCTATTTTTATGTCTAGGTTCCCCCTCCCCCTGTTAGGAGATGGGAGAGGTCACGA  
AAGAAAGGAATTTGTAGCCCTTGAGCCAGCCCGGGCCACAGAGCCTGCCACCAGAC  
AGGAAAAGCCCAGGGCTTACCAGCACAGGAGGAGCAAACCTCGCAGGCGAGCCTGG  
GTTGGCGCTGGTGGTCCCGGGTCGATGGCCCGCCATTCCCAGAAGCCGAGGCTATA  
GCTGCGTCACCTGCCCCGCCCTCCTCCCGAGTGAAGACCCCTAGAGGCTGAGCAGAC  
CCCAAAGGCGGTGCAATTCCATTGGCCCAAGGCAGAGGTGAGCGGCTGCTAATCCC  
CTCGGGAAGTGAAGGGACCCAGAGAGTCTGGTAGATGTGGGAGCTGGGGTTCAGGG  
CGAGACAGAGGGTGGGATGGGCAGAAGGGTCCAGGAAAGGAAAGTACTGGAGGGG  
AGTTGGGACAAAAGCAGCGACCAAGGGAACATCGCTTCAGTGACTGAAGCCAGGCA  
AAAGGAGCGGGAAGGATTATATGTAGCCTGGGACGCTTTCATAAACACTGATGACG  
TGTTTGTGCAAAGCAAGCAATTTGAGGAGAAACGCCTGGGACCTCGGAAAGAAGGA

bioRxiv preprint doi: <https://doi.org/10.1101/123456>; this version posted January 1, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.





ACAATTGCGAGGACCTAATTTAGGATCTCCAATAGCCACATAAAAAGCACAGCATG  
GCGGCAGACACCTGCAATTCCTGTCCCTGGAAGCACCTGTTGAGAATCCCAGAGACT  
CATTGGCCAAACACTCTATTCAATCAATGAAGTCCATATTCAGTGACAAAACCTTGAC  
TCAGAAACTAATGTGGAAAGCATCAGGAAGACAGCCAACATCTGGTCTCTACTCAT  
GCATGAATAAGGGATCCCAGAGAGAAGGGGAAGAAAAAGGAAGGAAGGAAGGAAG  
GAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGAGAGGGAGGAAAGGAGGG  
AGGGAAGGAAGGAAAGGGAAAGGAAAAAAGAGATGGGGAGGGAGGGAAGGAAAG  
GAAAGGGGGAGAAAGAAGAGAAGAAAGGAAAAATAAATAAATTTTCAGGGATTATT  
ACACCTTTAAATTTTATCCATAAAAGGTCATTTCCACCTGTTTGTCTGGAAGTAGAGT  
GGGATCCCTTATATAAGGGCAGTCTTTAACATAGTAGCATTTTATAAACCATTACAA  
ATTTTGAGTTTTCTCTACTTTTTATCCTCTACCATCTTCAAACCTGAAACTACAATTATT  
CCCACAAATGAAGAAAATGCTGTAAGAGTTTTACACACCCGAAGTGGGAAACTTAA  
GGATTAGACAAGTCTAACAATGAGAATGGGGAGAACAAAAAGAGACTGCACAGGG  
AGCCCTTTCTCTGCTTATAATCTTGACACTTGAGAAGCTAATTGACGCTGCATGACTA  
CTCAACTCTTTAAGCAAAACAATGCTGTTGTTTCATGAAAAGCACATAAAGTACATAT  
GTCCCATATATTCATCAAAATTTGCATGCAGCACATAATAGCAATCAAAGCAATAA  
CACCCACTGTTTCACAGAGACTTTAAACATGAAACTGGAACCTATGTCTAGTGTTTTGA  
CTTAGGGTACATAGTATGCTGTGTCTGTATGTACCAATGTTGATTTAGGTCATCAGA  
CAGCATTTGGAACATGTATCTTCAGGAGGAATCATTATGTATCCTGCATGAAATTC  
TCCACCTATGTTTATTCTCTTAGCCAGGTTTTTCTCTGATGGAGAAACATTGGGTTTG  
AGGTTTTACTCCCAGGTAACATTTAGGGAAAAGCTGTCTATGTTCTCAGTTTGGCTTT  
TATTTATGAGGGATGTTGGTATTCCAGAAAATTCTCTTTTGAAGAGATTACAATTTA  
GGTCAAAACAGAAAAATATGTAAAAAGTTATTGTTTTTATTAGTATTTTCATGTTCTTT  
TCTTTTTTAAAAATGGTATGCTTAGAACTAATTAAGATTAGATTAGATTAGATTAGA  
AAATAATCAGAGAGGGATTGATGAATGCTAAAGCATCATGAAAAATTCAAATTTT  
TTTGCTTCTAATTCAGAATCAATTAATTCATATTACTATAAAAGACAGCACGCCAG  
ATGTGTGCCAGCTGAGGAGTGGATAAACTGTGTAACGTGAGTGCTATGTAGAAACA  
GAAAGGAGTGAAGGGTTGATGTGCGCTGCAACATCTTGAAAACATTCGGCTACATG  
ATGGAAGCCAGGCACAAAAAGCCACATATTGCATGGTTATGTTTATATGAAATGTTT  
AAAATACATGGATTCTTAGCAAACAGAGTAAGATGTTACTTAGGGTCAGGAAAAGA  
TTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA  
TCCTAAGACGAAATAGTTGAGGTAGATATAGTTATATCCCTGTGGATATTGTAATAA  
ACCAGCATGCTGTGCTCTGAGAAGGGCCTAATGAAGGGGCAGGAGGAAGTGAAATG  
AGATGGTAGAAAGGAAAGTCATATACCATGGCTTCTCTCGTGGGTGGAATCTAGAT  
ATGTTAATATATTGACATAAAGGAAGGAATTGTTTAGGGAAGGATCAAACCAACA  
GGAGTGAGGGAGACAATAGGAACCAATGAGAGGCAAAGTTCATGGTCAATGTGTGT  
GGAGACACCATAATAAACTCCTTTTTTGTGTTGCTAACTAAAACCACTAAAATCTAA  
AAACAAAACATTTTTGCAACAAGAATTATTTATTATTCAATAAAGATGTTTAAATGGG  
GGAAGTTGAAGTTCATTGATAGTCTCATAAATCTTAAATGTATTAAACTGCTTTTTA  
CGTTTTTTATTATTAATTACTCTTGCTGTCATTATTATCATCATCATTATCGTCATCAT

CATCACTAATGCTTTTCACCATACACAAATGTAGGCAGAAGAGTGTAATCCACTTAG  
TGAGGCAATCTTGGAGAGGGAAAGGAAGCGGATGCGGGGCAGAGGCACACAGGAG  
GACAGTGAGAGGGGAAATGAACAAGAAAAAATGTGGACACATGCACAAAAATTCCA  
TAGTCCACTACATTACTTTGTATTCTAATATTAAGAAAATAATAAACCCATTTCTGTG  
CACTTATCACCCAGGCTCAACAGTTATCTTGGCCACAGATCCTGTCTCACTGCATCCT  
GTCCACCTGAGTCCACTTAGCGTTCTGAATCCAATCCAGGGCATGATGCTTACTCCT  
ACACAGAACTAAAGATTAAAGAGAGTTTAAAAGTAACCATGACATCTCTCTGTTTCCT  
TTAGCGATAAGTTCTTAATATTTATGGCTGCTTGTGTATGTTCTAATTTCTCTAATATT  
GTCACATTTAGTTGGCAACTACTTTGTTTGAATTGAGTTGGAGTTAAGGTCCCATAG  
GATTAATCTCAACATATTTCTATATTTATAAACTTTTCTCTCTTTGTGAAAGTTCCTTT  
GAGAAAACAAATATGCCCATATCTTTCTTTACAGGTCTTAAAAATGAGCTATAACAA  
AGTCCAAATAATTGAGAAGGATACTTTGTATGGACTCAGGAGCTTGACCCGGTTGCA  
CCTGGATCACAACAACATTGAGTTTATCAACCCCGAGGCGTTTTACGGACTCACCTT  
GCTCCGCTTGGTACATCTAGAAGGAAACCGGCTGACAAAGCTCCATCCAGACACAT  
TTGTCTCTTTGAGCTATCTCCAGATATTTAAAACCTCCTTCATTAAGNACCTGTACTT  
GTATGATAACTTCATTGACCTCCCTCCCAAAGAAATGGTCTCCTCTATGCCAAACC  
TAGAAAGCCTTTACTTGCATGGAAACCCATGGACCTGTGACTGCCATTTAAAGTGGT  
TGTCCGAGTGGATGCAGGGAAACCCAGGTAACCTATCTTGTTTGTGTTTCTTTTTTT  
ATARKACGTATTTTCCTCAATTTCAATTTAGAATGATATCCCAAAGTCCCCCATAACC  
TCCCCCCCCTTCCCTACCTACCCATTCCCATTTTTTTGGCCCTGGCATTCCCCTGTACT  
GGGGCATATAAAGTTTGCGTGTCCAATGGACCTCTCTTTCCAGTGATGGCCAACTAG  
GCCATCTTTTGATACATATGCAGCTAGAGTCAAGAGCTCTGGGGTACTGGTTAGTTC  
ATAATGTTGTTGCACCTACAGGGTTGAA (SEQ ID NO:20)



Figure 51

MPKRAHWGALS VVLLLLWGHPRVALACPHPCACYVPSEVHCTFRSLASVPAGIARHVE  
RINLGFNSIQALSETSFAGLTKLELLMIHGNEIPSIPDGALRDLSSLQVFKFSYNKLRVITG  
QTLQGLSNLMRLHIDHNKIEFIHPQAFNGLTSLRLLHLEGNLLHQLHPSTFSTFTFLDYFR  
LSTIRHLYLAENMVRTLPASMLRNMPLLENLYLQGNPWTCDCEMRWFLEWDAKSRGIL  
KCKKDKAYEGGQLCAMCFSPKKLYKHEIHKLKDMTCLKPSIESPLRQNRSRSEEEQEQ  
EEDGGSQ LILEKFQLPQWSISLNM TDEHGNMVNLVCDIKKPM DVYKIHLNQTDPPDIDIN  
ATVALDFECPMTRENYEKLWKLIA YYSEVPVKLHRELMLSKDPRVSYQYRQDADEEAL  
YYTG VRAQILAEP EWMQPSIDIQLNRRQSTAKKVLLSYTQYSQTISTKDTRQARGRS  
WVMIEPSGAVQRDQTVLEGGPCQLSCNVKASESPSIFWVLPDGSILKAPMDDPDSKFSIL  
SSGWLRIKSMEPSDSGLYQCIAQVRDEMDRMVYRVLVQSPSTQPAEKDTVITIGKNPGES  
VTLPCNALAIPEAHL SWILPNRRIINDLANTSHVYMLPNGTLSIPKVQVSDSGYYRCVAV  
NQQGADHFTVGITVTKKGSGLPSKRGRPGAKALSRVREDIVEDEGGSGMGDEENTSR  
RLHHPKDQEVFLKTKDDAINGDKKAKKGRRKLKLWKHSEKEPETNVAEGRRVFESRRR  
INMANKQINPERWADILAKVRGKNLPKGTEVPPLIKTTSPPSLSLEVTTPPPFAVSPPSASP  
VQTVTSAEESSADVPLLGE EHVLTGISSASMGLEHNHNGVILVEPEVTSTPLEEVVDDL  
SEKTEEITSTEGDLKGTAAPT LISEPYEPSPTLHTLDTVYEKP THEETATEGWSAADVGSS  
PEPTSSEYEPPLDAVSLAESEPMQYFDPDLET KSQPD EDKMKEDTFAHLTPTPTIWVND S  
STSQLFEDSTIGEPGVPGQSHLQGLTDNIHLVKSSLSTQDTLLIKKG MKEMSQTLQGGNM  
LEGDPTHSRSSSESEGQESKSITLPDSTLGIMSSMSPVKKPAETTVGTLLDKDTT TTTTPR  
QKVAPSSTMSTHPSRRRPNGRRLRP NKFRHRHKQTPPTTFAPSETFSTQPTQAPDIKISS  
QVESSLVPTAWVDNTVNTPKQLEMEKNAEPTSKGT PRRKHGKRPNKHYTPSTVSSRA  
SGSKPSPSPENKHRNIVTPSSETILLPRTVSLKTEGPYDSLDMYMTTTRKIYSSYPKVQETLP  
VTYKPTSDGKEIKDDVATNVDKHKSDILVTGESITNAIPTSRSLVSTMGEFKEESSPVGFP  
GTPTWNPSRTAQPGRLQTDIPVTTSGENLTD PPLLKELEDVDFTSEFLSSLTVSTPFHQEE  
AGSSTLSSIKVEVASSQAETTTLDQDHLETTVAILLSETRPQNHTPTAARMKEPASSPS  
TILMSLGQTTTTK PALPSPRISQASRDSKENVFLNYVGNPETEATPVNNEGTOHMSGPNE  
LSTPSSDRDAFNLSTKLELEKQVFGSRSLPRGPDSQRQDGRVHASHQLTRVPAKPILPTA  
TVRLPEMSTQSASRYFVTSQSPRHWTNKPEITTYPSGALPENKQFTTPRLSSTTIPLPLHM  
SKPSIPSKFTDRRTDQFNGYSKVFGNNNIPEARNPVGKPPSPRIPHY SNGRLPFFTNTLSF  
PQLGVTRRPQIPTSPAPVMRERKVIPGSYNRIHSHSTFHLDFGPPAPPLLHTPQT TGSPSTN  
LQNP MVSSSTQSSISFITSSVQSSGSFHQSSSKFFAGGPPASKFWSLGEKPQILTKSPQTVSV  
TAETD TVFPCEATGKPKPFVTWTKVSTGALMTPNTRIQRFEVLKNGTLVIRKVQVQDRG  
QYMCTASNLHGLDRMVLLSVTVQQPQILASHYQDVTVYLGDTIAMECLAKGTPAPQI  
SWIFPDRRVWQTVSPVESRITLHENRTLSIKEASFSDRGVYKCVASNAAGADSLAIRLHV  
AALPPVIHQEKLENISLPPGLSIHHCTAKAAPLPSVRWVLGDGTQIRPSQFLHGNLFVFP  
NGTLYIRNLAPKDSGRYECVAANLVGSARRTVQLNVQRAAANARITGTSPRRTD VRYG  
GTLKLD CSASGDPWPRILWRLPSKRMIDALFSFDSRIKVFANGTLV VKSVTDK



Figure 52

ATGCCCAAGC GCGCGCACTG GGGGGCCCTC TCCGTGGTGC TGATCCTGCT  
TTGGGGCCAT CCGCGAGTGG CGCTGGCCTG CCCGCATCCT TGTGCCTGCT  
ACGTCCCCAG CGAGGTCCAC TGCACGTTCC GATCCCTGGC TTCCGTGCCC  
GCTGGCATTG CTAGACACGT GGAAAGAATC AATTTGGGGT TTAATAGCAT  
ACAGGCCCTG TCAGAAACCT CATTTCAGG ACTGACCAAG TTGGAGCTAC  
TTATGATTCA CGGCAATGAG ATCCCAAGCA TCCCCGATGG AGCTTTAAGA  
GACCTCAGCT CTCTTCAGGT TTTCAAGTTC AGCTACAACA AGCTGAGAGT  
GATCACAGGA CAGACCCTCC AGGGTCTCTC TAACTTAATG AGGCTGCACA  
TTGACCACAA CAAGATCGAG TTTATCCACC CTCAAGCTTT CAACGGCTTA  
ACGTCTCTGA GGCTACTCCA TTTGGAAGGA AATCTCCTCC ACCAGCTGCA  
CCCCAGCACC TTCTCCACGT TCACATTTTT GGATTATTTT AGACTCTCCA  
CCATAAGGCA CCTCTACTTA GCAGAGAACA TGGTTAGAAC TCTTCCTGCC  
AGCATGCTTC GGAACATGCC GCTTCTGGAG AATCTTTACT TGCAGGGAAA  
TCCGTGGACC TGCATTGTG AGATGAGATG GTTTTTGGAA TGGGATGCAA  
AATCCAGAGG AATTCTGAAG TGTA AAAAGG ACAAAGCTTA TGAAGGCGGT  
CAGTTGTGTG CAATGTGCTT CAGTCCAAAG AAGTTGTACA AACATGAGAT  
ACACAAGCTG AAGGACATGA CTTGTCTGAA GCCTTCAATA GAGTCCCCTC  
TGAGACAGAA CAGGAGCAGG AGTATTGAGG AGGAGCAAGA ACAGGAAGAG  
GATGGTGGCA GCCAGCTCAT CCTGGAGAAA TTCCAACCTGC CCCAGTGGAG  
CATCTCTTTG AATATGACCG ACGAGCACGG GAACATGGTG AACTTGGTCT  
GTGACATCAA GAAACCAATG GATGTGTACA AGATTCACTT GAACCAAACG  
GATCCTCCAG ATATTGACAT AAATGCAACA GTTGCCCTTG ACTTTGAGTG  
TCCAATGACC CGAGAAAACCT ATGAAAAGCT ATGGAAATTG ATAGCATACT  
ACAGTGAAGT TCCCGTGAAG CTACACAGAG AGCTCATGCT CAGCAAAGAC  
CCCAGAGTCA GCTACCAGTA CAGGCAGGAT GCTGATGAGG AAGCTCTTTA  
CTACACAGGT GTGAGAGCCC AGATTCTTGC AGAACCAGAA TGGGTCATGC  
AGCCATCCAT AGATATCCAG CTGAACCGAC GTCAGAGTAC GGCCAAGAAG  
GTGCTACTTT CCTACTACAC CCAGTATTCT CAAACAATAT CCACCAAAGA  
TACAAGGCAG GCTCGGGGCA GAAGCTGGGT AATGATTGAG CCTAGTGGAG  
CTGTGCAAAG AGATCAGACT GTCCTGGAAG GGGGTCCATG CCAGTTGAGC  
TGCAACGTGA AAGCTTCTGA GAGTCCATCT ATCTTCTGGG TGCTTCCAGA  
TGGCTCCATC CTGAAAGCGC CCATGGATGA CCCAGACAGC AAGTTCTCCA  
TTCTCAGCAG TGGCTGGCTG AGGATCAAGT CCATGGAGCC ATCTGACTCA  
GGCTTGTACC AGTGCATTGC TCAAGTGAGG GATGAAATGG ACCGCATGGT  
ATATAGGGTA CTTGTGCAGT CTCCCTCCAC TCAGCCAGCC GAGAAAGACA  
CAGTGACAAT TGGCAAGAAC CCAGGGGAGT CGGTGACATT GCCTTGCAAT  
GCTTTAGCAA TACCCGAAGC CCACCTTAGC TGGATTCTTC CAAACAGAAG  
GATAATTAAT GATTTGGCTA ACACATCACA TGTATACATG TTGCCAAATG  
GAACTCTTTC CATCCCAAAG GTCCAAGTCA GTGATAGTGG

TTACTACACAGA TGTGTGGCTG TCAACGAGCA AGGGGCAGAC CATTTTACGG  
TGGGAATCAC AGTGACCAAG AAAGGGTCTG GCTTGCCATC CAAAAGAGGC  
AGACGCCCAG GTGCAAAGGC TCTTTCCAGA GTCAGAGAAG ACATCGTGGA  
GGATGAAGGG GGCTCGGGCA TGGGAGATGA AGAGAACACT TCAAGGAGAC  
TTCTGCATCC AAAGGACCAA GAGGTGTTCC TCAAAACAAA GGATGATGCC  
ATCAATGGAG ACAAGAAAGC CAAGAAAGGG AGAAGAAAGC TGAAACTCTG  
GAAGCATTCTG GAAAAAGAAC CAGAGACCAA TGTTGCAGAA GGTCGCAGAG  
TGTTTGAATC TAGACGAAGG ATAAACATGG CAAACAAACA GATTAATCCG  
GAGCGCTGGG CTGATATTTT AGCCAAAGTC CGTGGGAAAA ATCTCCCTAA  
GGGCACAGAA GTACCCCCGT TGATTAAAC CACAAGTCCT CCATCCTTGA  
GCCTAGAAGT CACACCACCT TTTCTGCTG TTTCTCCCC CTCAGCATCT  
CCTGTGCAGA CAGTAACCAG TGCTGAAGAA TCCTCAGCAG ATGTACCTCT  
ACTTGGTGAA GAAGAGCACG TTTTGGGTAC CATTTCTCA GCCAGCATGG  
GGCTAGAACA CAACCACAAT GGAGTTATTC TTGTTGAACC TGAAGTAACA  
AGCACACCTC TGGAGGAAGT TGTTGATGAC CTTTCTGAGA AGACTGAGGA  
GATAACTTCC ACTGAAGGAG ACCTGAAGGG GACAGCAGCC CCTACACTTA  
TATCTGAGCC TTATGAACCA TCTCCTACTC TGCACACATT AGACACAGTC  
TATGAAAAGC CCACCCATGA AGAGACGGCA ACAGAGGGTT GGTCTGCAGC  
AGATGTTGGA TCGTCACCAG AGCCACATC CAGTGAGTAT GAGCCTCCAT  
TGGATGCTGT CTCCTTGGCT GAGTCTGAGC CCATGCAATA CTTTGACCCA  
GATTTGGAGA CTAAGTCACA ACCAGATGAG GATAAGATGA AAGAAGACAC  
CTTTGCACAC CTTACTCCAA CCCCCACCAT CTGGGTTAAT GACTCCAGTA  
CATCACAGTT ATTTGAGGAT TCTACTATAG GGGAACCAGG TGTCCCAGGC  
CAATCACATC TACAAGGACT GACAGACAAC ATCCACCTTG TGAAAAGTAG  
TCTAAGCACT CAAGACACCT TACTGATTAA AAAGGGTATG AAAGAGATGT  
CTCAGACACT ACAGGGAGGA AATATGCTAG AGGGAGACCC CACACACTCC  
AGAAGTTCTG AGAGTGAGGG CCAAGAGAGC AAATCCATCA CTTTGCCTGA  
CTCCACACTG GGTATAATGA GCAGTATGTC TCCAGTTAAG AAGCCTGCGG  
AAACCACAGT TGGTACCCTC CTAGACAAAG ACACCACAAC AGTAACAACA  
ACACCAAGGC AAAAAGTTGC TCCGTCATCC ACCATGAGCA CTCACCCTTC  
TCGAAGGAGA CCCAACGGGA GAAGGAGATT ACGCCCCAAC AAATTCCGCC  
ACCGGCACAA GCAAACCCCA CCCACAACCT TTGCCCCATC AGAGACTTTT  
TCTACTCAAC CAACTCAAGC ACCTGACATT AAGATTTCOA GTCAAGTGGA  
GAGTTCTCTG GTTCCTACAG CTTGGGTGGA TAACACAGTT AATACCCCCA  
AACAGTTGGA AATGGAGAAG AATGCAGAAC CCACATCCAA GGGAACACCA  
CGGAGAAAAC ACGGGAAGAG GCCAAACAAA CATCGATATA CCCCTTCTAC  
AGTGAGCTCA AGAGCGTCCG GATCCAAGCC CAGCCCTTCT CCAGAAAATA  
AACATAGAAA CATTGTTACT CCCAGTTCAG AAATAATACT TTTGCCTAGA  
ACTGTTTCTC TGAAAACCTGA GGGCCCTTAT GATTCCTTAG ATTACATGAC  
AACCACCAGA AAAATATATT CATCTTACCC TAAAGTCCAA GAGACACTTC  
CAGTCACATA

TAAACCCACA TCAGATGGAA AAGAAATTAA GGATGATGTT GCCACAAATG  
TTGACAAACA TAAAAGTGAC ATTTTAGTCA CTGGTGAATC AATTACTAAT  
GCCATACCAA CTTCTCGCTC CTTGGTCTCC ACTATGGGAG AATTTAAGGA  
AGAATCCTCT CCTGTAGGCT TTCCAGGAAC TCCAACCTGG AATCCCTCAA  
GGACGGCCCA GCCTGGGAGG CTACAGACAG ACATACCTGT TACCACTTCT  
GGGGAAAATC TTACAGACCC TCCCCTTCTT AAAGAGCTTG AGGATGTGGA  
TTTCACTTCC GAGTTTTTGT CCTCTTTGAC AGTCTCCACA CCATTTTACC  
AGGAAGAAGC TGGTTCTTCC ACAACTCTCT CAAGCATAAA AGTGGAGGTG  
GCTTCAAGTC AGGCAGAAAC CACCACCCTT GATCAAGATC ATCTTGAAAC  
CACTGTGGCT ATTCTCCTTT CTGAAACTAG ACCACAGAAT CACACCCCTA  
CTGCTGCCCC GATGAAGGAG CCAGCATCCT CGTCCCCATC CACAATTCTC  
ATGTCTTTGG GACAAACCAC CACCACTAAG CCAGCACTTC CCAGTCCAAG  
AATATCTCAA GCATCTAGAG ATTCCAAGGA AAATGTTTTT TTGAATTATG  
TGGGGAATCC AGAAACAGAA GCAACCCAG TCAACAATGA AGGAACACAG  
CATATGTCAG GGCCAAATGA ATTATCAACA CCCTCTTCCG ACCGGGATGC  
ATTTAACTTG TCTACAAAGC TGGAATTGGA AAAGCAAGTA TTTGGTAGTA  
GGAGTCTACC ACGTGGCCCA GATAGCCAAC GCCAGGATGG AAGAGTTCAT  
GCTTCTCATC AACTAACCAG AGTCCCTGCC AAACCCATCC TACCAACAGC  
AACAGTGAGG CTACCTGAAA TGTCCACACA AAGCGCTTCC AGATACTTTG  
TAACTTCCCA GTCACCTCGT CACTGGACCA ACAAACCGGA AATAACTACA  
TATCCTTCTG GGGCTTTGCC AGAGAACAAA CAGTTTACAA CTCCAAGATT  
ATCAAGTACA ACAATTCTC TCCATTGCA CATGTCCAAA CCCAGCATTC  
CTAGTAAGTT TACTGACCGA AGAACTGACC AATTCAATGG TTAATCCAAA  
GTGTTTGGA ATAACAACAT CCCTGAGGCA AGAAACCCAG TTGGAAGCC  
TCCCAGTCCA AGAATTCTC ATTATTCCAA TGGAAGACTC CCTTTCTTTA  
CCAACAAGAC TCTTTCTTTT CCACAGTTGG GAGTCACCCG GAGACCCAG  
ATACCCACTT CTCCTGCCCC AGTAATGAGA GAGAGAAAAG TTATTCCAGG  
TTCTTACAAC AGGATACATT CCCATAGCAC CTTCCATCTG GACTTTGGCC  
CTCCGGCACC TCCGTTGTTG CACACTCCGC AGACCACGGG ATCACCTCA  
ACTAATTAC AGAATATCCC TATGGTCTCT TCCACCCAGA GTTCTATCTC  
CTTTATAACA TCTTCTGTCC AGTCCTCAGG AAGCTTCCAC CAGAGCAGCT  
CAAAGTTCTT TGCAGGAGGA CCTCCTGCAT CCAAATTCTG GTCTCTTGGG  
GAAAAGCCCC AAATCCTCAC CAAGTCCCCA CAGACTGTGT CCGTCACCGC  
TGAGACAGAC ACTGTGTTCC CCTGTGAGGC AACAGGAAAA CCAAAGCCTT  
TCGTTACTTG GACAAAGGTT TCCACAGGAG CTCTTATGAC TCCGAATACC  
AGGATACAAC GGTTTGAGGT TCTCAAGAAC GGTACCTTAG TGATACGGAA  
GGTTCAAGTA CAAGATCGAG GCCAGTATAT GTGCACCGCC AGCAACCTGC  
ACGGCCTGGA CAGGATGGTG GTCTTGCTTT CGGTCACCGT GCAGCAACCT  
CAAATCCTAG CCTCCCACTA CCAGGACGTC ACTGTCTACC TGGGAGACAC  
CATTGCAATG GAGTGTCTGG CCAAAGGGAC CCCAGCCCCC CAAATTTCTT  
GGATCTTCCC TGACAGGAGG GTGTGGCAAA CTGTGTCCCC

1000  
900  
800  
700  
600  
500  
400  
300  
200  
100  
0

CGTGGAGAGC CGCATCACCC TGCACGAAAA CCGGACCCTT TCCATCAAGG  
AGGCGTCCTT CTCAGACAGA GCGGTCTATA AGTGCGTGGC CAGCAATGCA  
GCCGGGGCGG ACAGCCTGGC CATCCGCCTG CACGTGGCGG CACTGCCCCC  
CGTTATCCAC CAGGAGAAGC TGGAGAACAT CTCGCTGCCC CCGGGGCTCA  
GCATTCACAT TCACTGCACT GCCAAGGCTG CGCCCCTGCC CAGCGTGCGC  
TGGGTGCTCG GGGACGGTAC CCAGATCCGC CCCTCGCAGT TCCTCCACGG  
GAACTTGTTT GTTTTCCCCA ACGGGACGCT CTACATCCGC AACCTCGCGC  
CCAAGGACAG CGGGCGCTAT GAGTGCGTGG CCGCCAACCT GGTAGGCTCC  
GCGCGCAGGA CGGTGCAGCT GAACGTGCAG CGTGCAGCAG CCAACGCGCG  
CATCACGGGC ACCTCCCCGC GGAGGACGGA CGTCAGGTAC GGAGGAACCC  
TCAAGCTGGA CTGCAGCGCC TCGGGGGACC CCTGGCCGCG CATCCTCTGG  
AGGCTGCCGT CCAAGAGGAT GATCGACGCG CTCTTCAGTT TTGATAGCAG  
AATCAAGGTG TTTGCCAATG GGACCCTGGT GGTGAAATCA GTGACGGACA  
AAGATGCCGG AGATTACCTG TGCGTAGCTC GAAATAAGGT TGGTGATGAC  
TACGTGGTGC TCAAAGTGGA TGTGGTGATG AAACCGGCCA AGATTGAACA  
CAAGGAGGAG AACGACCACA AAGTCTTCTA CGGGGGTGAC CTGAAAGTGG  
ACTGTGTGGC CACCGGGCTT CCAATCCCG AGATCTCCTG GAGCCTCCCA  
GACGGGAGTC TGGTGAAGTC CTTTCATGCAG TCGGATGACA GCGGTGGACG  
CACCAAGCGC TATGTCGTCT TCAACAATGG GACACTCTAC TTAAACGAAG  
TGGGGATGAG GGAGGAAGGA GACTACACCT GCTTTGCTGA AAATCAGGTC  
GGGAAGGACG AGATGAGAGT CAGAGTCAAG GTGGTGACAG CGCCCGCCAC  
CATCCGGAAC AAGACTTACT TGGCGGTTCA GGTGCCCTAT GGAGACGTGG  
TCACTGTAGC CTGTGAGGCC AAAGGAGAAC CCATGCCCAA GGTGACTTGG  
TTGTCCCCAA CCAACAAGGT GATCCCCACC TCCTCTGAGA AGTATCAGAT  
ATACCAAGAT GGCCTCTCC TTATTCAGAA AGCCAGCGT TCTGACAGCG  
GCAACTACAC CTGCCTGGTC AGGAACAGCG CGGGAGAGGA TAGGAAGACG  
GTGTGGATTG ACGTCAACGT CCAGCCACCC AAGATCAACG GTAACCCCAA  
CCCCATCACC ACCGTGCGGG AGATAGCAGC CGGGGGCAGT CGGAACTGA  
TTGACTGCAA AGCTGAAGGC ATCCCCACCC CGAGGGTGTT ATGGGCTTTT  
CCCGAGGGTG TGGTTCTGCC AGCTCCATAC TATGGAAACC GGATCACTGT  
CCATGGCAAC GGTTCCTGG ACATCAGGAG TTTGAGGAAG AGCGACTCCG  
TCCAGCTGGT ATGCATGGCA CGCAACGAGG GAGGGGAGGC GAGGTTGATC  
GTGCAGCTCA CTGTCCTGGA GCCCATGGAG AAACCCATCT TCCACGACCC  
GATCAGCGAG AAGATCACGG CCATGGCGGG CCACACCATC AGCCTCAACT  
GCTCTGCCGC GGGGACCCCG ACACCCAGCC TGGTGTGGGT CCTTCCCAAT  
GGCACCGATC TGCAGAGTGG ACAGCAGCTG CAGCGCTTCT ACCACAAGGC  
TGACGGCATG CTACACATTA GCGGTCTCTC CTCGGTGGAC GCTGGGGCCT  
ACCGCTGCGT GGCCCGCAAT GCCGCTGGCC ACACGGAGAG GCTGGTCTCC  
CTGAAGGTGG GACTGAAGCC AGAAGCAAAC AAGCAGTATC ATAACCTGGT  
CAGCATCATC AATGGTGAGA CCCTGAAGCT CCCCTGCACC CCTCCCGGGG  
CTGGGCAGGG ACGTTTCTCC TGGACGCTCC CCAATGGCAT GCATCTGGAG



**Figure 53**

ATGAAGGTAAGGCAAGGAGGAAATCACCTGCTTGTCTGGTCTCCTTTGCTGTGATCTGCCTGGTTCGCCACC  
CCTGGGGGCAAGGCCCTGTCTCGCCGCTGTGCCTGTTATATGCCTACGGAGGTACACTGCACATTTTCGG  
TACCTGACTTCCATCCCAGACAGCATCCCGCCCAATGTGGAACGCATCAATTTAGGATACAACAGCTTG  
GTTAGATTGATGGAACAGATTTTTCTGGCCTGACCAAACCTGGAGTTACTCATGCTTCACAGCAATGGC  
ATTCACACAATCCCTGACAAGACCTTCTCAGATTTGCAGGCCCTGCAGGTCTTAAAAATGAGCTATAAT  
AAAGTCCGAAAACCTCAGAAAGATACTTTTTATGGCCTCAGGAGCTTGACACGATTGCACATGGACCAC  
AACAATATTGAGTTTATAAAACCAGAGGTTTTTTATGGGCTCAACTTTCTCCGCTGGTGCACCTTGGAA  
GGAAATCAGCTCAGTAAAGCTCCACCCAGATACATTTGTCTCTTTGAGCTACCTCCAGATATTTAAATC  
TCTTTTCATTAAGTTCCCTATACTTGTCTGATAACTTCCTGACCTCCCTCCCTCAAGAGATGGTCTCCTAT  
ATGCCTGACCTAGACAGCCTTTACCTGCATGGAACCCATGGACCTGTGATTGCCATTTAAAGTGGTTG  
TCTGACTGGATACAGGAGAGCCAGATGTAATAAAATGCAAAAAGATAGAAGTCCCTCTAGTGCTCAG  
CAGTGTCCACTTTGTCATGAACCTTAGGACTTCTAAAGGCAAGCCGTTAGCTATGGTCTCAGCTGCAGCT  
TTCCAGTGTGCCAAGCCAACCATTTGACTCATCCCTGAAATCAAAGAGCCTGACTATTCTGGAAGACAGT  
AGTTCTGCTTTTCATCTCTCCCAAGGTTTCATGGCACCCCTTTGGCTCCCTCACTTTGAATATGACAGAT  
CAGTCTGGAAATGAAGCTAACATGGTCTGCAGTATTCAAAGCCCTCAAGGACATCACCCATTGCATTTC  
ACTGAAGAAAATGACTACATCGTGCTAAATACTTTCATTTTCAACATTTTGGTGTGCAACATAGATTAC  
GGTCACATTCAGCCAGTGTGGCAAATTTTGGCTTTGTACAGTGATTCTCCTCTGATACTAGAAAGGAGC  
CACTTGCTTAGTGAAACACCGCAGCTCTATTACAAATATAAACAGGTGGCTCCTAAGCCTGAAGACATT  
TTTACCAACATAGAGGCAGATCTCAGAGCAGATCCCTCTTGGTTAATGCAAGACCAAATTTCCCTTGACAG  
CTGAACAGAACTGCCACCACATTTCAGTACATTACAGATCCAGTACTCCAGTGATGCTCAAATCACTTTA  
CCAAGAGCAGAGATGAGGCCAGTGAAACACAAATGGACTATGATTTCAAGGGATAACAATACTAAGCTG  
GAACATACTGTCTTGGTAGGTGGAACCGTTGGCCTGAACTGCCAGGCCAAGGAGACCCACCCACAC  
GTGGATTGGCTTCTAGCTGATGGAAGTAAAGTGAGAGCCCTTATGTGAGTGGAGTGGACGGATCCTA  
ATAGACAAAAGTGGAAAATTTGGAATCCAGATGGCTGATAGTTTGGACACAGGCGTATATCACTGTATA  
AGCAGCAATTATGATGATGCAGATATTCTCACCTATAGGATAACTGTGGTAGAACCTTTGGTCCGAGCC  
TATCAGGAAAATGGGATTTCATCACACAGTTTTTCATTGGTGAACACTTGATCTTCCATGCCATTCTACT  
GGTATCCCAGATGCCCTCTATTAGCTGGGTATTCCAGGAAACAATGTGCTCTATCAGTCATCAAGAGAC  
AAGAAAGTTCTAAACAATGGCACATTAAGAATATTACAGGTCACCCCGAAAGACCAAGGTTATTATCGC  
TGTGTGGCAGCCAACCCATCAGGGGTTGATTTTTTGGATTTTCCAAGTTTCAGTCAAGATGAAAGGACAA  
AGGCCCTTGGAGCATGATGGAGAAACAGAGGATCTGGACTTGATGAGTCCAATCCTATTGCTCATCTT  
AAGGAGCCACCAGGTGCACAACCTCCGTACATCTGCTCTGATGGAGGCTGAGGTTGGAACACACCTCA  
AGCACAAAGTAAGAGGCACAACCTATCGGGAATTAACACTCCAGCGAGCTGGAGATTCAACACATCCAGCT  
TTTAGGGAGAATAGGAGGCATTTCCCTCCCTCTGCTAGGAGAATTGACCCACAACATTGGGCGGCACTG  
TTGGAGAAAGCTAAAAAGATGCTATGCCAGACAAGCGAGAAAATACCACAGTGAGCCACCCCAAGTG  
GTCACCCAACTCCCAAACATACCTGGTGAAGAAGACGATTCTCAGGCATGCTCGCTCTACATGAGGAA  
TTTATGGTCCCGGCCACTAAAGCTTTGAACCTTCCAGCAAGGACAGTGACTGCTGACTCCAGAACAATA  
TCTGATAGTCTTATGACAAACATAAATTATGGCACAGAAATCTCTCCTGTTGTGAATTACAAAATACTA  
CCACCTGAAGAACCCACAGATTTCAAACCTGTCTACTGCTATTAAAACACAGCCATGTCAAAGAATATA  
AACCCAAACATGTCAAGCCAAATACAAGGCACAACCAATCAACATTTCATCCACTGTCTTTCCACTGCTA  
CTTGGAGCAACTGAATTTTCAGGACTCTGACCAGATGGGAAGAGGAAGAGAGCATTTCCAAAGTAGACCC  
CCAATAACAGTAAGGACTATGATCAAAGATGTCAATGTCAAATGCTTAGTAGCACCACCAACAACTA  
TTATTAGAGTCAGTAAATACCACAAATAGTCATCAGACATCTGTAAGAGAAGTGAGTGAACCCAGGCAC  
AATCACTTCTATTCTCAGACTACTCAAATACTTAGCACCTCCACGTTCCCTTCAGATCCACACACAGCT  
GCTCATTCTCAGTTTCCGATCCCTAGAAATAGTACAGTTAACATCCCGCTGTTTCAGACGCTTTGGGAGG  
CAGAGGAAAATTTGGCGGAAGGGGGCGGATTATCAGCCCATATAGAAGTCCAGTTCTGCGACGGCATAGA  
TACAGCATTTTTCAGGTCAACAACACAGAGGTTCTTCTGAAAAAGCACTACTGCATTCTCAGCCACAGTG  
CTCAATGTGACATGTCTGTCTTCTCCAGGGAGAGGCTCACCACTGCCACAGCAGCATTTGTCTTTT  
CCAAGTGCTGCTCCCATCACCTTCCCCAAAGCTGACATTGCTAGAGTCCCATCAGAAGAGTCTACAAC  
CTAGTCCAGAATCCACTATTACTACTTGAGAACAAACCCAGTGTAAGAGAAAACAACACCCACAAATAAAA  
TATTTTCAGGACTGAAATTTCCCAAGTGACTCCAACCTGGTGCAGTCATGACATATGCTCCAACATCCATA  
CCCATGGAAAAAACTCACAAAGTAAACGCCAGTTACCCACGTTGTGTCTAGCACCAATGAAGCTAAAAGA  
GATTTCAGTGATTACATCGTCACTTTTCAGGTGCTATCACCAAGCCACCAATGACTATTATAGCCATTACA  
AGGTTTTCAAGAAGGAAAAATTCCTGGCAACAGAACTTTGTAAATAACCATAACCEAAAAGGCAGATTA  
AGGAATCAACATAAAGTTAGTTTACAAAAAGCACAGCTGTGATGCTTCTTAAACATCTCTGCTTTA  
CCACAGAGACAAAGTTCCCTTTCCATTTCAACACACTTTCAACAAGTGTGATGCAAAATTCATCTAAT  
ACCTTGACTACCGCTCACCACTACGACCAAAACACACAATCCTGGAAGTCTTCCAACAAGAAAGGAG  
CTTCCCTTCCACCCCTTAACCTATGCTTCTCTAGTATTATAAGCAAAGACTCAAGTACAAAAAGCATC  
ATATCAACGCAACAGCAATACCAGCAACAACCTACCTTCCCTGCATCTGTCTATCACTTTATGAAACC  
CAAACAGAGAGATCTAGAGCACAAACAATACAAAGAGAACAGGAGCCTCAAAAGAAGAACAGGACTGAC  
CCAAACATCTCTCCAGACCAGAGTTCTGGCTTCACTACACCCACTGCTATGACACCTCCTGCTCTGGCA



T T C A C T C A T T C C C C A C C A G A A A C A C A A C T G G G A T T T C A A G C A C A A T C A G T T T T C A T T C A A G A A C T C T T  
A A T C T G A C A G A T G T G A T T G A A G A A C T A G C C C A A G C A A G T A C T C A G A C T T T G A A G A G C A C A A T T G C T T C T  
G A A C A A C T T T G T C C A G C A A A T C A C A C C A G A G T A C C A C A A C T A G G A A A G C A T C A T T A G A C A C T C C C A T A  
C C A C C A T T C T T G A G C A G C A G T G C T A C T C T A A T G C C A G T T C C C A T C T C C C C T C C C T T T A C T C A G A G A C A  
G T T A C T G A C A C A C G T G G C G A C T C C C A T T T C C G G C T T A T G A C A A A T A C A G T G G T C A A G C T G C A C G A A T C C  
T C A A G G C A C A A T C T C C A A A T G C C A A G T T C A C A A T T G G A A C C A C T C A C T T C A T C T A C C T C T A A T C T G T T A  
C A T T C T A C T C C C A T G C C A G C A C T A A C A C A G T T A A A T C A C A G A A T T C C A A A T T A A C T C C A T C T C C C T G G  
G C A G A A T A C C A A T T T T G G C A C A A A C C A T A C T C A G A C A T T G C T G A A A A A G G C A A A A G C C A G A A G T A A G C  
A T G T T G G C T A C T A C A G G C C T G T C C G A G G C C A C C A C T C T T G T T C A G A T T G G G A T G G A C A G A A G A A C A C A  
A A G A A G A G T G A C T T T G T A T A A G A A A C C A G T T C A A G A A G C A A C A A C T T C C A A A C T C C T T C C C T T T G A C T C T  
T T G T C T A G G T A T A T A T T T G A A A A G C C C A G G A T A G T T G G A G G A A A G C T G C A A G T T T T A C T A T T C C A G C T  
A A C T C A G A T G C C T T T C T T C C C T G T G A A G C T G T T G G A A A T C C C C T G C C C A C C A T T C A T T G G A C C A G A G T T  
T C A G G A C T T G A T T T A T C T A G A G G A A A C C A G A A T A G C A G G G T C C A G G T T C T C C C C A A T G G T A C C C T G T C C  
A T C C A G A G G G T G G A A A T T C A G G A C C G C G G A C A G T A C T T G T G T T C C G C A T C C A A T C T G T T T G G C A C A G A C  
C A C C T T C A T G T C A C C T T G T C T G T G G T T T C C T A T C C T C C C A G G A T C C T G G A G A G A C G T A C C A A A G A G A T C  
A C A G T T C A T T C C G G A A G C A C T G T G G A A C T G A A G T G C A G A G C A G A A G G T A G G C C A A G C C C T A C A G T T A C C  
T G G A T T C T T G C A A A C C A A C A G T T G T C T C A G A A T C A T C C C A G G G A A G T A G G C A G G C T G T G G T G A C G G T T  
G A C G G A A C A T T G G T C C T C C A C A A T C T C A G T A T T T A T G A C C G T G G C T T T T A C A A A T G T G T G C C A G C A A C  
C C A G G T G G C C A G G A T T C A C T G C T G G T T A A A A T A C A A G T C A T T G C A G C A C C A C C T G T T A T T C T A G A G C A A  
A G G A G G C A A G T C A T T G T A G G C A C T T G G G G T G A A G T T T A A A A C T G C C C T G T A C T G C A A A A G G A A C T C C T  
C A G C C C A G C G T T T A C T G G G T C C T C T G A T G G C A C T G A A G T G A A A C C A T T A C A G T T T A C C A A T T C C A A G  
T T G T T C T T A T T T T C A A A T G G G A C T T T G T A T A T A A G A A A C C T A G C C T C T T C A G A C A G G G C A C T T A T G A A  
T G C A T T G C T A C C A G T T C C A C T G G T T C G G A G C G A A G A G T A G T A A T G C T T A C A A T G G A A G A G C G A G T G A C C  
A G C C C C A G G A T A G A A G C T G C A T C C C A G A A A A G G A C T G A A G T G A A T T T T G G G G A C A A A T T A C T A C T G A A C  
T G C T C A G C C A C T G G G G A G C C C A A A C C C C A A A T A A T G T G G A G G T T A C C A T C C A A G G C T G T G G T C G A C C A G  
T G G A G C T G G A T C C A C G T C T A C C C T A A T G G A T C C C T G T T T A T T G G A T C A G T A A C A G A A A A A G A C A G T G G T  
G T C T A C T T G T G T G T G G C A A G A A A C A A A A T G G G G G A T G A T C T G A T A C T G A T G C A T G T T A G C C T A A G A C T G  
A A A C C T G C C A A A A T T G A C C A C A A G C A G T A T T T T A G A A A G C A A G T G C T C C A T G G G A A A G A T T T C C A A G T A  
G A T T G C A A A G C T T C C G G C T C C C C A G T G C C A G A G A T A T C T T G G A G T T T G C C T G A T G G A C C A T G A T C A A C  
A A T G C A A T G C A A G C C G A T G A C A G T G G C C A G A G A T T A T A C T T G C T A T G C C C A G A A C C C T A G G G A A A  
T A C T T C A A C A A A G T T G G G G T A G C G G A G G A A G G A G A T T A T A C T T G C T A T G C C C A G A A C C C T A G G G A A A  
G A T G A A A T G A A G T C C A C T T A A C A G T T A T A A C A G C T G C T C C C C G G A T A A G G C A G A G T A A C A A A C C A A C  
A A G A G A A T C A A A G C T G G A G A C A C A G C T G T C C T T G A C T G T G A G G T C A C T G G G G A T C C C A A A C C A A A A A T A  
T T T T G G T T G C T G C C T T C C A A T G A C A T G A T T T C C T T C C A T T G A T A G G T A C A C A T T T C A T G C C A A T G G G  
T C T T T G A C C A T C A A C A A A G T G A A A C T G C T C G A T T C T G G A G A G T A C G T A T G T G T A G C C C G A A A T C C C A G T  
G G G G A T G A C A C C A A A A T G T A C A A A C T G G A T G T G G T C T A A A C C T C C A T T A A T C A A T G G T C T G T A T A C A  
A A C A G A A C T G T T A T T A A A G C C A C A G C T G T G A G A C A T T C C A A A A A C A C T T T G A C T G C A G A G C T G A A G G G  
A C A C C A T C T C C T G A A G T C A T G T G G A T C A T G C C A G A C A A T A T T T T C C T C A C A G C C C C A T A C T A T G G A A G C  
A G A A T C A C A G T C C A T A A A A A T G G A A C C T T G G A A A T T A G G A A T G T G A G G C T T T C A G A T T C A G C C G A C T T T  
A T C T G T G T G G C C C G A A A T G A A G G T G G A G A G A G C G T G T T G G T A G T A C A G T T A G A A G T A C T G G A A A T G C T G  
A G A A G A C C G A C A T T T A G A A A T C C A T T T A A T G A A A A A T A G T T G C C C A G C T G G G A A A G T C C A C A G C A T T G  
A A T T G C T C T G T T G A T G G T A A C C C A C C A C T G A A A T A A T C T G G A T T T T A C C A A T G G C A C A C G A T T T T C C  
A A T G G A C C A A A A G T T A T C A G T A T C T G A T A G C A A G C A A T G G T T C T T T T A T C A T T T C T A A A A C A A C T C G G  
G A G G A T G C A G G A A A A T A T C G C T G T G C A G C T A G G A A T A A A G T T G G C T A T A T T G A G A A A T A G T C A T A T T A  
G A A A T T G G C C A G A A G C C A G T T A T T C T T A C C T A T G C A C C A G G G A C A G T A A A A G G C A T C A G T G G A G A A T C T  
C T A T C A C T G C A T T G T G T G T C T G A T G G A A T C C C T A A G C C A A A T A T C A A A T G G A C T A T G C C A A G T G G T T A T  
G T A G T A G A C A G G C C T C A A A T T A A T G G G A A A T A C A T A T T G C A T G A C A A T G G C A C C T A G T C A T T A A A G A A  
G C A A C A G C T T A T G A C A G A G G A A A C T A T A T C T G T A A G G C T C A A A A T A G T G T T G G T C A T A C A C T G A T T A C T  
G T T C C A G T A A T G A T T G T A G C C T A C C C T C C C C G A A T T A C A A A T C G T C C A C C C A G G A G T A T T G T C A C C A G G  
A C A G G G G C A G C C T T T C A G C T C C A C T G T G T G G C C T T G G G A G T T C C C A A G C C A G A A A T C A C A T G G G A G A T G  
C C T G A C C A C T C C C T T C T C T C A A C G G C A A G T A A A G A G A G G A C A C A T G G A A G T G A G C A G C T T C A C T T A C A A  
G G T A C C C T A G T C A T T C A G A A T C C C C A A C C T C C G A T T C T G G G A T A T A C A A A T G C A C A G C A A A G A A C C C A  
C T T G G T A G T G A T T A T G C A G C A A C G T A T A T T C A A G T A A T C T G A

**Figure 54**

ORIGIN

```

1      MKVKGRGITC LLVSFAVICL VATPGGKACP RRCACYPMTE VHCTFRYLTS
51     IPDSIPPNVE RINLGYNLSV RLMETDFSGL TKLELLMLHS NGIHTIPDKT
101    FSDLQALQVL KMSYNKVRKL QKDTFYGLRS LTRLHMDHNN IEFINPEVFY
151    GLNFLRLVHL EGNQLTKLHP DTFVSLSYLQ IFKISFIKFL YLSDNFLTSL
201    PQEMVSYMPD LDSLYLHGPN WTCDCHLKWL SDWIOEKPDV IKCKKDRSPS
251    SAQQCPLCMN PRTSKGKPLA MVSAAAFQCA KPTIDSSLKS KSLTILEDSS
301    SAFISPOGFM APFGSLTLNM TDQSGNEANM VCSIQKPSRT SPIAFTEEND
351    YIVLNTSFST FLVCNIDYGH IQPVWQILAL YSDSPLILER SHLLSETPQL
401    YYKYKQVAPK PEDIFTNIEA DLRADPSWLM ODQISLQLNR TATTFSTLQI
451    QYSSDAQITL PRAEMRPVKH KWTMISRDNN TKLEHTVLVG GTVGLNCPGQ
501    GDPTPHVDWL LADGSKVRAP YVSEDGRILI DKSGKLELQM ADSFDTGVYH
551    CISSNYDDAD ILTYRITVVE PLVEAYQENG IHHTVFIGET LDLPCHSTGI
601    PDASISWVIP GNNVLYQSSR DKKVLNNGTL RILQVTPKDQ GYYRCVAANP
651    SGVDFLIFQV SVKMKGQRPL EHDGETEGSG LDESNPIAHL KEPPGAQLRT
701    SALMEAEVGK HTSSTSQRHN YRELTQRRG DSTHRRFREN RRHFPPSARR
751    IDPQHWAAALL EKAKKNAMPD KRENTTVSPP PVVTQLPNIP GEEDDSSGML
801    ALHEEFMVPA TKALNLPART VTADSRTISD SPMTNINYGT EFSPVVSQI
851    LPPEEPTDFK LSTAIKTTAM SKNINPTMSS QIQGTTNQHS STVFPLLLGA
901    TEFQSDSQMG RGREHFQSRP PITVRTMIKD VNVKMLSSTT NKLLLESVNT
951    TNSHQTSVRE VSEPRHNHFY SHTTQILSTS TFPSPDHTAA HSQFPIPRNS
1001   TVNIPLFRFF GRQRKIGGRG RIISPYRTPV LRRHRSIFR STTRGSSEKS
1051   TTAFSATVLN VTCLSLCLPE RLTTATAALS FPSAAPITFP KADIARVPSE
1101   ESTTLVQNPL LLLLENKPSVE KTTPTIKYFR TEISQVTPTG AVMTYAPTSI
1151   PMEKTHKVNA SYPRVSSTNE AKRDSVITSS LSGAITKPPM TIIAITRFSR
1201   RKIPWQONFV NNHNPKGRLR NQHKVSLQKS TAVMLPKTSP ALPQRQSSPF
1251   HFTTLSTSVM QIPSNTLTTA HHTTTKTHNP GSLPTKKELP FPPLNPMPLPS
1301   IISKDSSTKS IISTQTAIPA TTPTFPASVI TYETQTERS AQTIOREQEP
1351   QKKNRDTPNI SPDQSSGFTT PTAMTPPALA FTHSPPENTT GISSTISFHS
1401   RTLNLTDVIE ELAQASTQTL KSTIASETTL SSKSHQSTTT RKASLDTPIP
1451   PFLSSSATLM PVPISPPFTQ RAVTDTRGDS HFRLMTNTVV KLHESSRHNL
1501   QMPSSQLEPL TSSTSNNLLHS TMPALTTVK SQNSKLTPSP WAEYQFWHKP
1551   YSDIAEKGKK PEVSMLATTG LSEATTLVSD WDGQKNKKKS DFDKPKVQEA
1601   TTSKLLPFDS LSRYIFEKPR IVGGKAASFT IPANSDAFLP CEAVGNPLPT
1651   IHWTRVSGLD LSRGNQNSRV QVLPNGTSLI QRVEIQDRGQ YLCSASNLFQ
1701   TDHLHVTLVS VSYPPRILER RTKEITVHSG STVELKCRAE GRPSPTVTWI
1751   LANQTVVSES SQGSRQAVVT VDGTLLVHLN SIYDRGFYKC VASNPGGQDS
1801   LLVKIQVIAA PPVILEQRRQ VIVGTWGESL KLPCTAKGTP QPSVYVWVLS
1851   GTEVKPLQFT NSKLFLFSNG TLYIRNLASS DRGTYECIAT SSTGSERRVV
1901   MLTMEERVTS PRIEASQKR TEVNFQDKLL LNCSATGEPK PQIMWRLPSK
1951   AVVDQWSWIH VYPNGSLFIG SVTEKDSGVY LCVARNMKMG DLILMHVSLR
2001   LKPAKIDHKQ YFRKQVLHGK DFQVDCASG SPVPEISWSL PDGTMINNAM
2051   QADDSGHRTR RYTLFNNGTL YFNKVGVAEE GDYTCYAQNT LGKDEMKVHL
2101   TVITAAPRIR QSNKTNKRIK AGDTAVLDCE VTGDPKPKIF WLLPSNDMIS
2151   FSIDRYTFHA NGSLTINKVK LLDSGEYVCV ARNPSGDDTK MYKLDVVSKP
2201   PLINGLYTNR TVIKATAVRH SKKHFDCAE GTPSPEVMWI MPDNIFLTAP
2251   YYGSRITVHK NGTLEIRNVR LSDSADFICV ARNEGGSVL VVQLEVLEML
2301   RRPTFRNPFN EKIVAQLGKS TALNCSVDGN PPPEIIWILP NGTRFSNGPQ
2351   SYQYLIASNG SFIISKTTRE DAGKYRCAAR NKVGYIEKLV ILEIGQKPMI
2401   LTYAPGTVKG ISGESLSLHC VSDGIPKPN I KWTMPSGYV DRPQINGKYI
2451   LHDNGTLVIK EATAYDRGNY ICKAQNSVGH TLITVPMIV AYPPIRITNR
2501   PRSIVTRTGA AFQLHCVALG VPKPEITWEM PDHSLSTAS KERTHGSEQ
2551   HLQGTLLVIQ PQTSDSGIYK CTAKNPLGSD YAATYIQVI*

```



Figure 56

## pCM-H-608-663-N-term

